PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

PCT

To: JANE MASSEY LICATA
LAW OFFICES OF JANE MASSEY LICATA
66 E. MAIN STREET
MARLTON, NJ 08053

Docket System ✓
Status Report ✓
Docket Book ——

NP = 11/3/98

NOTIFICATION OF TRANSMITTAL OF INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Rule 71.1)

Date of Mailing *(day/month/year)*	08 SEP 1998
Applicant's or agent's file reference: BDA-0029	IMPORTANT NOTIFICATION

International application No.	International filing date *(day/month/year)*	Priority Date *(day/month/year)*
PCT/US97/07491	02 MAY 1997	03 MAY 1996

Applicant
CUBICCIOTTI, ROGER S.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. REMINDER

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices)(Article 39(1))(see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	NEIL LEVY aco
Facsimile No. (703) 305-3230	Telephone No. (703) 308-1235

Form PCT/IPEA/416 (July 1992)★

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and Administrative Instructions, Section 422)

To:

LICATA, Jane, Massey
Law Offices of Jane Massey Licata
Woodland Falls Corporate Park
Suite 201
210 Lake Drive East
Cherry Hill, NJ 08002
ETATS-UNIS D'AMERIQUE

Date of mailing (day/month/year)
02 July 1997 (02.07.97)

Applicant's or agent's file reference
BDA-0029

IMPORTANT NOTIFICATION

International application No.
PCT/US97/07491

International filing date (day/month/year)
02 May 1997 (02.05.97)

1. The following indications appeared on record concerning:

[X] the applicant [X] the inventor [ ] the agent [ ] the common representative

Name and Address
CUBICCIOTTI, Roger, S.
46 Church Street
Montclair, NJ 07042
United States of America

State of Nationality: US
State of Residence: US
Telephone No.:
Facsimile No.:
Teleprinter No.:

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

[ ] the person [ ] the name [X] the address [ ] the nationality [ ] the residence

Name and Address
CUBICCIOTTI, Roger, S.
258 Midland Avenue
Montclair, NJ 07042
United States of America

State of Nationality: US
State of Residence: US
Telephone No.:
Facsimile No.:
Teleprinter No.:

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

[X] the receiving Office
[X] the International Searching Authority
[ ] the International Preliminary Examining Authority
[ ] the designated Offices concerned
[ ] the elected Offices concerned
[ ] other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer
M. Abidine

Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

United States Patent and Trademark Office
(Box PCT)
Crystal Plaza 2
Washington, DC 20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 13 January 1998 (13.01.98)	
International application No. PCT/US97/07491	**Applicant's or agent's file reference** BDA-0029
International filing date (day/month/year) 02 May 1997 (02.05.97)	**Priority date** (day/month/year) 03 May 1996 (03.05.96)
Applicant CUBICCIOTTI, Roger, S.	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

25 November 1997 (25.11.97)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland**	Authorized officer M. Abidine
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference BDA-0029	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/US97/07491	International filing date *(day/month/year)* 02 MAY 1997	Priority date *(day/month/year)* 03 MAY 1996
International Patent Classification (IPC) or national classification and IPC IPC(6): A61K 9/50 and US Cl.: 424/450		
Applicant CUBICCIOTTI, ROGER S.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 3 sheets.

   ☐ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of — sheets.

3. This report contains indications relating to the following items:

   I ☒ Basis of the report

   II ☐ Priority

   III ☐ Non-establishment of report with regard to novelty, inventive step or industrial applicability

   IV ☐ Lack of unity of invention

   V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

   VI ☐ Certain documents cited

   VII ☐ Certain defects in the international application

   VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
25 NOVEMBER 1998	25 JULY 1998
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer neil levy Telephone No. (703) 308-1235

Form PCT/IPEA/409 (cover sheet) (January 1994)*

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US97/07491

I. Basis of the report

1. This report has been drawn on the basis of *(Substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments)*:

[X] the international application as originally filed.

[X] the description, pages 1-21 , as originally filed.
pages NONE , filed with the demand.
pages NONE , filed with the letter of _____.
pages _____ , filed with the letter of _____.

[X] the claims, Nos. 1-8 , as originally filed.
Nos. NONE , as amended under Article 19.
Nos. NONE , filed with the demand.
Nos. NONE , filed with the letter of _____.
Nos. _____ , filed with the letter of _____.

[X] the drawings, sheets/~~fig~~ NONE , as originally filed.
sheets/~~fig~~ NONE , filed with the demand.
sheets/~~fig~~ NONE , filed with the letter of _____.
sheets/~~fig~~ _____ , filed with the letter of _____.

2. The amendments have resulted in the cancellation of:

[X] the description, pages NONE.

[X] the claims, Nos. NONE.

[X] the drawings, sheets/~~fig~~ NONE.

3. [ ] This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the ~~Supplemental Box~~ Additional observations below (Rule 70.2(c)).

4. Additional observations, if necessary:

NONE

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US97/07491

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. STATEMENT

Novelty (N)	Claims	1-8	YES
	Claims	none	NO
Inventive Step (IS)	Claims	1-8	YES
	Claims	none	NO
Industrial Applicability (IA)	Claims	1-8	YES
	Claims	NONE	NO

2. CITATIONS AND EXPLANATIONS

Claims 1-8 meet the criteria for Industrial Applicability under PCT Article 33(4) as providing useful drug compositions for disease treatment.

Claims 1-8 meet the criteria set out in PCT Article 33(2)-(4), because the prior art does not teach or fairly suggest the synthetic receptor-prodrug complex.

---------- NEW CITATIONS ----------

NONE

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 9/50

A1

(11) International Publication Number: WO 97/41840

(43) International Publication Date: 13 November 1997 (13.11.97)

(21) International Application Number: PCT/US97/07491

(22) International Filing Date: 2 May 1997 (02.05.97)

(30) Priority Data:
60/016,566 3 May 1996 (03.05.96) US

(71)(72) Applicant and Inventor: CUBICCIOTTI, Roger, S. [US/US]; 258 Midland Avenue, Montclair, NJ 07042 (US).

(74) Agents: LICATA, Jane, Massey et al.; Law Offices of Jane Massey Licata, 66 E. Main Street, Marlton, NJ 08053 (US).

(81) Designated States: AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GE, GH, HU, IL, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: PRODRUG COMPOSITIONS AND DRUG DELIVERY METHODS USING SYNTHETIC RECEPTORS

(57) Abstract

The present invention provides prodrug and multi-prodrug complexes comprising drugs specifically bound to synthetic receptors in such a manner that active drug becomes available only in the presence of a targeted pathophysiologic receptor. Methods for preparation and use of these prodrug complexes in drug delivery systems are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/07491

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 9/50
US CL :424/450

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/450, 484, 485, 490, 1.45, 1.57, 9.34

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

NONE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	EP 0496520 A1 (MERCK & CO., INC.) 29 July 1992, pages 1, 5-12.	1, 2, 5, 6, --- 1-8
X	US 4,193,983 A (ULLMAN et al) 18 March 1980, columns 7-10, 14, 19.	1-8
X -- Y	US 4,931,553 A (GILL et al) 05 June 1990, columns 2-4, 12.	1-8 --- 1-8
X -- Y	US 5,490,991 A (ENRIQUEZ et al) 13 February 1996, column 5, lines 48-55, column 6, lines 3-7, 20 - line 18 of column 7, examples 3-5.	1, 2, 5, 6 --- 1-8

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
18 JUNE 1997

Date of mailing of the international search report
04 AUG 1997

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
Neil Levy

Telephone No. (703) 305-9646

Form PCT/ISA/210 (second sheet)(July 1992)★

JOURNAL OF
MEDICINAL
CHEMISTRY

Volume 37, Number 9 April 29, 1994

Perspective

Applications of Combinatorial Technologies to Drug Discovery. 1. Background and Peptide Combinatorial Libraries

Mark A. Gallop, Ronald W. Barrett, William J. Dower, Stephen P. A. Fodor, and Eric M. Gordon*

Affymax Research Institute, 4001 Miranda Avenue, Palo Alto, California 94304

Received December 21, 1993

Biographies

Ronald W. Barrett received his Ph.D. in Pharmacology from Rutgers University in 1983 and did postdoctoral work at the Addiction Research Foundation in Palo Alto in the field of opiate receptors. From 1986 to 1989, he worked in the Neuroscience Research Area at Abbott Laboratories in Abbott Park, IL. He joined Affymax in 1989 where he established the receptor pharmacology group. Dr. Barrett is presently Vice President and Director of Receptor Pharmacology.

William J. Dower received his Ph.D. in Biology from the University of California, San Diego, and did postdoctoral work at Stanford University where he studied steroid control of specific mRNA metabolism and the amplification of genes mediated by chemotherapeutic agents. He joined BioRad Laboratories in 1984 where as a member of the newly formed molecular biology unit, he introduced the electroporation technique for the high-efficiency transformation of bacteria. He joined Affymax in 1989 to establish a molecular biology research group. Dr. Dower is Director of Molecular Biology at Affymax.

Stephen P. A. Fodor received his Ph.D. in Chemistry at Princeton University and was an NIH postdoctoral fellow in Chemistry at the University of California, Berkeley. He joined Affymax in 1989 as a Staff Scientist in Optical Technologies and became Director of Physical Sciences. Dr. Fodor's group led the development of new technologies, merging photolithography with combinatorial solid-phase chemistry. In 1993 he joined Affymetrix as Scientific Director where he is currently using oligonucleotide arrays to study a variety of DNA molecular recognition processes.

Mark A. Gallop received his Ph.D. from the University of Cambridge, England, for research in the area of organo-transition metal chemistry and did postdoctoral work in the laboratories of Peter G. Schultz at the University of California, Berkeley. He joined Affymax in 1990 and is currently a Senior Scientist and leader of the Combinatorial Chemistry Group.

Eric M. Gordon received his Ph.D. in 1973 from the University of Wisconsin—Madison and did postdoctoral work at Yale University. His research interests include the rational design of enzyme inhibitors, the chemistry of amino acids, peptides, and natural products, and combinatorial chemistry. Dr. Gordon has authored 150 papers and U.S. patents in these areas. From 1974 to 1992 he worked at the Squibb Institute for Medical Research (subsequently the Bristol Myers Squibb Pharmaceutical Research Institute) in Princeton, NJ, most recently as a Director of Medicinal Chemistry. In 1992, he joined the Affymax Research Institute where he is currently Vice President of Research and Director of Chemistry.

A. Introduction

Recent trends in the search for novel pharmacological agents have focused on the preparation of "chemical libraries" as potential sources of new leads for drug discovery. Chemical libraries are intentionally created collections of differing molecules which can be prepared either synthetically or biosynthetically and screened for biological activity in a variety of different formats (e.g., libraries of soluble molecules; libraries of compounds tethered to resin beads, silica chips, or other solid supports; recombinant peptide libraries on bacteriophage and other biological display vectors; etc). Combinatorial chemistry is a type of synthetic strategy which leads to large chemical libraries. For the purposes of this discussion, combinatorial chemistry may be defined as the systematic and repetitive, covalent connection of a set of different "building blocks" of varying structures to each other to yield a large array of diverse molecular entities.

Traditionally, new medicinal chemical lead structures have originated from the isolation of natural products from microbiological fermentations, plant extracts, and animal sources; from screening of pharmaceutical company compound databases; and more recently through the application of both mechanism-based and structure-based approaches to rational drug design.[1,2] The advent of high throughput, automated techniques has made possible the robotized screening of in excess of hundreds of thousands

of individual compounds per year, per drug target. The availability of this capability, combined with a major, worldwide shift in emphasis by the drug industry toward more cost-effective pharmaceutical products, has exacerbated the need for a continuous flow of huge numbers of novel molecules. Several studies have shown that the average cost of creating a new molecular enitity in a major pharmaceutical company is around $7500/compound.[3] A significant portion of this figure relates to an intrinsic feature of contemporary medicinal chemistry which has, by and large, involved the "hand-crafted" serial synthesis and testing of individual molecular entities. A major theme of the industrial revolution was replacement of highly expensive, individualized manufacturing with the concept of mass production to greatly decrease manpower and cost. Though this period began in America more than a century ago, to date the idea has not been significantly applied to the preparation of pharmaceutical leads.

To address inefficiencies inherent in the contemporary new lead discovery process, researchers have recently turned to the concept of using combinatorial chemical technologies and devising novel assay formats for rapidly evaluating these compound collections.[4-7] Thus by employing a building-block collection (*vide infra*), and systematically assembling these blocks in many combinations using chemical, biological, or biosynthetic procedures, it is possible to create "chemical libraries" as vast populations of molecules. This approach has clear parallels in Nature, which eons ago applied the building block approach for creation of oligonucleotides, carbohydrates, and peptides/proteins, through the combination of nucleosides, sugars, and amino acids respectively.

An essential starting point for the generation of molecular diversity is an assortment of small, reactive molecules which may be considered chemical building blocks. The universe of structural diversity accessible through assembly of even a small set of building-block elements is potentially large, and unleashing the power inherent in the building-block approach is crucial to the success of the combinatorial method. The building-block argument is easily illustrated, and its implications are profound. Theoretically, the number of possible different individual compounds, N, prepared by an ideal combinatorial synthesis is determined by two factors; the number of blocks available for each step "b", and the number of synthetic steps in the reaction scheme, x. If an equal number of building blocks are used in each reaction step, then $N = b^x$. If the number of building blocks for each step varies (e.g., b, c, d in a three-step synthesis), then $N = bcd$. Exploitation of a basis set of (for example) 100 interchangeable building blocks permits the theoretical synthesis of 100 million tetrameric or 10 billion pentameric chemical entities (see Figure 1).

In this perspective we will review the technologies and strategies that have recently emerged for using combinatorial chemical methods in ligand discovery. Historically, these methods made their first appearance in the development of peptide libraries. The ready availability of a large and structurally diverse range of amino acid building blocks, a highly refined, generic coupling chemistry, and the fact that small peptides are biologically and pharmaceutically key molecules, focused early efforts on peptide chemistry as a useful vehicle for exploring the power and conceptual issues attached to combinatorial ligand discovery. In addition, biological systems have been

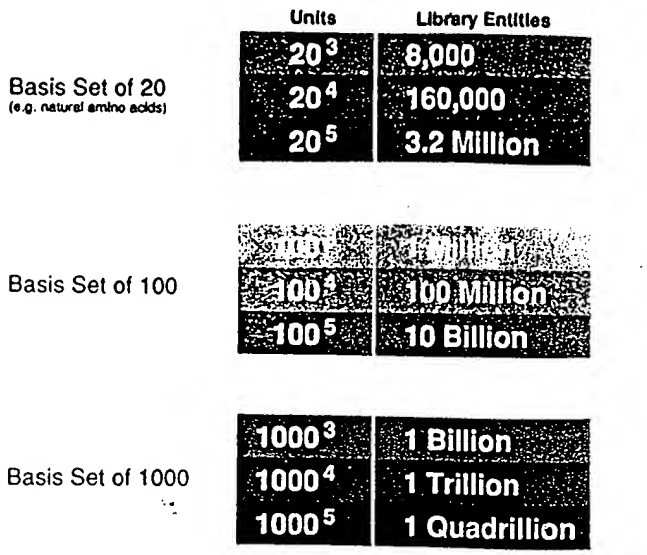


Figure 1. Creating chemical diversity from a basis set of building blocks.

used to generate vast collections of peptides through the expression of randomly synthesized DNA. The ability to selectively enrich individual peptide sequences from these recombinant libraries has demonstrated the power and versatility of peptide diversity for *de novo* ligand identification. Generalization of the combinatorial strategy has led to construction of collections of other natural polymers (e.g., oligonucleotides) and unnatural (synthetic) polymeric libraries. Currently there is an acute interest in the generation of small-molecule libraries, through "combinatorial organic synthesis (COS)": the rapid synthesis of enormous numbers of diverse, low molecular weight, nonpolymeric organic molecules. In surveying progress in the field to date, we will also attempt to define criteria and constraints which must be imposed on the creation of libraries useful for drug discovery.

An essential element of the combinatorial discovery process is that *one must be able to extract the information made available by library screening*. Put another way, creating large quantities of molecular diversity for ligand discovery is insufficient, unless there is a format at hand to capture the information, which in this case is the chemical structures of active compounds. In the course of this overview, particular attention will be directed to the assay methodologies employed in screening combinatorial libraries and to the interrelationship between these methods, the diversity display formats, and the mechanisms for determining the structure of selected molecules that emerge from library evaluation. Two divergent embodiments of the combinatorial discovery exercise will be highlighted: (i) *random screening*, where the task is to identify a lead compound in the absence of any structural information about active molecules, and (ii) *directed screening* or *chemical analoging*, where the objective is to evaluate closely related structural analogs of a lead molecule, establish SAR, and optimize biological potency. These activities raise different issues in terms of both the synthetic combinatorial strategies and assay parameters that must be applied.

The above cursory analysis, at face value, makes clear that one or a few combinatorial experiments is sufficient to create more different molecules than exist in the combined compound files of the worldwide pharmaceutical industry! Recent research suggests that medicinal chem-

istry and drug discovery may be entering a new era in which vast numbers of small molecules may be readily accessible by combinatorial approaches. The power to synthetically create and evaluate huge numbers of known and future pharmacophores is unprecedented and suggests that combinatorial technologies may rapidly intersect and ultimately shortcut the traditional path of medicinal chemistry. The generation and screening of such immense numbers of compounds will demand innovative tools for data handling and analysis. Questions such as what should be the composition of a building-block collection suitable for drug discovery, what kinds of new synthetic and screening strategies are suitable, what formats for diversity creation are relevant, how to integrate both diversity generation and screening, and ultimately how to develop a new paradigm to dramatically shorten the drug-discovery process will be considered and discussed.

B. Biological Approaches to Generating Molecular Diversity

The harnessing of biological systems for the generation of peptide diversity mimics the evolutionary creation of protein diversity; however, artificial evolution is greatly accelerated by the introduction of diversity into the system at a much higher rate than occurs naturally. The source of diversity is the combinatorial chemical synthesis of mixtures of oligonucleotides. Oligonucleotide synthesis is a well-characterized chemistry that allows tight control of the composition of the mixtures created. The degenerate sequences produced are then placed into an appropriate genetic context for expression as peptides.

The particular context of expression of the randomly-encoded peptides is a key feature in the practical application to screening. Because a very large number of different peptides can be produced by biological expression, a highly efficient means of selecting the most interesting of these peptides is required. Depending on the type of activity that is sought, a variety of selection schemes are possible. If, for example, a particular catalytic activity is desired, a genetic screen might be devised that would endow organisms that express that activity with a growth advantage. The application emphasized in this review is the identification of new ligands, an approach based on the affinity purification of active peptides by adsorption on a target receptor. With each of the biological systems that will be described, identification of individual active peptides is enabled by physical linkage of the peptides to their encoding DNA. Capturing the peptides also recovers the DNA, and the replicative nature of the biological system is employed to amplify and sequence the DNA, thereby identifying the active peptides.

In practice, the process begins with the synthesis of single-stranded DNA. Synthetic oligonucleotides have been used to create genetic mutants since the 1970s. Initially, precise changes were made in regulatory and protein-encoding regions in order to study the phenotypes of the resulting mutants.[8,9] Later, pools of mutants were produced by introducing a variety of base changes into sites in the genome. It was soon realized that large collections, or libraries, of mutant clones could be acquired by inserting fully random cassettes of oligonucleotides into the targeted loci.[10-13] Because of the combinatorial nature of random cassette mutagenesis, the number of mutant strains in these libraries becomes much too large to analyze individually, and population selection schemes of the sort mentioned above are required.

Conventional expression cloning techniques rely on the detection of the binding activity of an expressed protein that is associated geographically or physically with its encoding DNA. The identification of individual clones that bind to a particular probe can be detected among thousands to millions of inert clones by the use of phage- or colony-lift techniques. Here, the clones are grown at rather high density on agar surfaces and "lifted" onto polymeric membranes to adsorb the macromolecules from the clones in a pattern that mirrors the arrangement of the colonies on the growth plates. Detection is achieved with a binding protein, such as an antibody, that is appropriately labeled. Once binding is detected on the membrane, the corresponding colony or plaque on the growth plate can be located and isolated for further analysis.[14,15] A method that has been employed in the cloning of receptors invokes the surface expression of clone libraries on eukaryotic cells, followed by selection of the desired clones through affinity adsorption of cells displaying the receptor on a matrix of immobilized ligand. This selection method has been dubbed "panning".[16,17] Upon isolation of the cells, the cloned fragments of DNA they contain are rescued and sequenced. These methods are commonly used, powerful means of isolating clones from cDNA and genomic libraries of normal size (up to $\sim 10^7$ clones), but they become impractical with the very large random libraries of 10^8–10^9 clones typically employed in ligand discovery projects. Recently, the advent of extraordinarily efficient cloning tools has permitted the construction of clone libraries of up to 10^{11} members.[18]

Because the number of different peptides one can create by this combinatorial approach is astronomical, and because the expectation is that peptides with the appropriate structural characteristics to serve as ligands for a given receptor will be rare, the need for methods capable of conveniently screening billions of clones is apparent. Several strategies for selecting very rare peptide ligands have been described. These are categorized as peptide display techniques and are distinguished by the peptide presentation format and by the nature of the linkage between the peptide-encoding DNA and the peptide itself. This association effectively tags the peptide for identification during screening, since in practice the peptide sequence is revealed by sequencing of the DNA.

Several approaches for displaying peptides or proteins on the surface of microorganisms have been developed. Each of these methods employs the fusion of the "guest" peptides to a cell surface protein of the host. A number of reports describe the use of the maltose receptor of *E. coli* (the outer membrane protein, LamB) as a peptide fusion partner.[19] Oligonucleotides have been inserted into plasmids encoding the LamB gene to produce peptides fused into one of the extracellular loops of the protein. These peptides are available for binding to antibodies and can elicit an immune response when the cells are administered to animals. Other cell surface proteins, e.g., OmpA,[20] PhoE,[21] and PAL,[22] as well as large bacterial surface structures have served as vehicles for peptide display.[†] Peptides have been fused to pilin, a protein which polymerizes to form the pilus—a conduit for interbacterial exchange of genetic information.[23] Because of its role in

[†] OmpA, outer membrane protein A; PhoE, phosphate-limitation inducible outer membrane protein; PAL, peptidoglycan-associated lipoprotein.

interacting with other cells, the pilus provides a useful support for the presentation of peptides to the extracellular environment. Another large surface structure pressed into service for peptide display is the bacterial motive organ, the flagellum. Fusion of peptides to the subunit protein flagellin offers a dense array of many peptide copies on the host cells.[24] Surface proteins of other bacterial species have also served as peptide fusion partners. Examples include the *Staphylococcus* protein A and the outer membrane protease IgA of *Neisseria*.[25,26]

Peptides and Proteins Displayed on Phage Particles. Systems for presenting peptides on bacterial cells have potential as random library vectors for affinity selection of ligands, but to date, only the LamB system has been used in this way.[27] An alternative system, which presents peptides on the much less complex surface of bacteriophage particles, has found wide acceptance. These vectors are derived from a family of filamentous phages of *Escherichia coli* that include phage m13, f1, and fd. In 1985, Smith published a seminal paper describing the display of peptides on filamentous phage particles.[28] The minor coat protein pIII was chosen as the fusion partner, and a fragment of the restriction enzyme EcoRI was placed in the middle of pIII. The guest peptide could be detected on the surface of mature phage with anti-EcoRI antibodies. The normal function of pIII is to mediate adsorption to the host cell (via the sex pilus) as a prelude to the entry of the phage into the bacterial cell. pIII is the largest protein of the virion (~40 kD) and is well suited by its natural function to the display of peptides for binding to exogenous proteins. The C-terminus of the pIII protein is embedded in the phage coat while the N-terminus projects into the environment. In 1988, Parmley and Smith extended this work,[29] relocating the fused peptides to the N-terminal domain of pIII, an insertion site less disruptive to the normal function of the protein. This report presented a detailed method for selecting phage bearing peptide ligands and demonstrated the isolation of a few phage carrying fragments of β-galactosidase from an enormous excess of inert phage by affinity purification on an anti-β-galactosidase antibody. Although the authors proposed the method primarily as a tool for cDNA expression library cloning, they also suggested that a library of short, randomly created peptides could serve as an "epitope library" for mapping the binding specificities of antibodies.

To realize the potential of such random libraries required the construction of libraries of extraordinary size. In 1990 three groups, each utilizing a variant of filamentous phage as a pIII–peptide expression vector, reported the creation of libraries of $\sim 10^7$ to $>10^8$ recombinants, and the successful selection from these libraries of specific ligands for monoclonal antibodies,[30,31] and for the biotin-binding protein, streptavidin.[32] The much more abundant major coat protein, pVIII, of the filamentous phage can also serve as a peptide expression vehicle, as demonstrated initially by Il'chev.[33] Greenwood *et al.*[34] cloned oligonucleotides into the single pVIII gene of the fd genome, leading to fusion of the expressed peptide to all of the several thousand copies of pVIII on each phage particle. This format was effective in expressing peptides up to six residues in length. However, peptides longer than six residues impeded the assembly of the virions, greatly reducing the yield of phage produced. This limitation was addressed by utilizing a phagemid that contained a pVIII gene into which the oligonucleotide was inserted. Infection of cells carrying these phagemids with a "helper" phage provided all of the phage functions including the wild-type pVIII. Adjustment of the level of expression of the peptide-fused pVIII permitted expression of the recombinant peptide on a minority of the coat molecules and allowed efficient assembly of phage carrying peptides of a range of sizes. The phagemid gene specifying the pVIII–peptide fusion protein was placed under control of the inducible tac promotor, and the authors estimated the fusion protein to be incorporated into mature virions as 10% of the total pVIII for one 12-mer peptide and as 30% for another unrelated 12-mer.[34] Expression levels of decapeptide–pVIII fusions driven by an arabinose promotor are 8–25% of total pVIII, as measured by peptide sequencing analysis.[35] Felici *et al.*[36] employed a pVIII phagemid vector to prepare a library of $\sim 10^8$ clones displaying random peptides nine residues in length. This library was screened on an antibody raised against a fragment of IL-1β, and the recovered peptides contained a three-residue subset of the immunogenic peptide.

Phage display methods have found wide use in the engineering of proteins and protein domains. Although the focus in this review is the use of display techniques for small peptide discovery, a brief discussion of their role in protein engineering is relevant. The availability of recombinant display technology is revolutionizing the field of antibody engineering. Since the first reports of antibody domains displayed on the major[37] or minor[38-40] coat proteins of phage, the field has progressed very rapidly. cDNA libraries of single-chain Fvs and assembled Fabs representing the repertoire of heavy and light chains of an animal are cloned into a phage vector to create many possible combinations of the associated chains, potentially creating new antigen specificities. The libraries may be derived from animals immunized with the antigen of interest or may even be made from a "naive" source not previously exposed to the target antigen.[40,41] The antibody library is then selected against an immobilized antigen, often yielding a large number of reactive clones. The antibodies recovered in the early rounds of selection are usually of modest affinities, with K_d values in the range of high nanomolar to mid micromolar. The affinities can be improved using several strategies, collectively referred to as "*in vitro* affinity maturation"—a process that mimics the selection of high-affinity antibodies that occurs *in vivo* during the course of an immune response. *In vitro* affinity improvement is accomplished by simply continuing selection of the pool of antibodies on the immobilized antigen, or by introducing sequence variation into the enriched antibody pool and reapplying selection. The variation is produced by shuffling heavy and light chains to create new chain pairings and by mutagenizing the complementarity determining regions (CDRs) to create new variations of the combining site with improved complementarity to the antigen. Among the benefits of this burgeoning area are the rapid development of antibodies with customized specificities and high affinities, and most importantly, the development of completely human antibodies, some with specificity for even human antigens, and with the potential for application as therapeutic agents. Excellent reviews of this fast-moving field are available.[42,43]

The engineering of other classes of proteins has also benefited from the application of recombinant display technology. Human growth hormone (hGH) was the fir-

example of a foreign protein displayed on filamentous phage in a functionally active form with respect to binding its receptor.[44] Lowman *et al.*[45] extended this work, accomplishing the rapid artificial evolution of the wild-type hGH to a much more potent form via phagemid display selection. Markland and co-workers[46] expressed functional bovine pancreatic trypsin inhibitor (BPTI) on the major coat protein of phage particles. This was achieved by introducing a second pVIII gene, controlled by an inducible promotor, into the genome of phage m13. The level of expression of the BPTI-pVIII fusion was estimated to be 30–60 molecules per phage. This group also reported constructing a small library of BPTI mutants and screening it against human neutrophil elastase (HNE) to find the most potent inhibitor of HNE yet described.[47] Pannekoek *et al.*[48] expressed another protease inhibitor, human plasminogen-activator inhibitor 1 (PAI-1) on pIII, and showed retention of the ability to bind t-PA. A library of lightly mutagenized variants of PAI-1 was constructed for the purpose of probing structure–activity relationships of the inhibitor molecules. The enzymes, alkaline phosphatase and trypsin, have been expressed in catalytically active forms on phage surfaces,[49,50] opening the way for applied evolution of altered catalytic properties. The means to produce huge collections of mutants and to rapidly select for those with enhanced binding and other features is destined to have an enormous impact on the field of protein engineering.

As profound as developments in the area of protein engineering may be, the greatest significance of recombinant display methods for medicinal chemistry is the promise of small peptides as lead compounds for drug development. Phage-borne peptide libraries have been screened against a variety of receptors. Since the initial reports identifying families of peptide ligands for antibodies appeared, a number of antibodies have been mapped for their epitope specificities.[51] Those with continuous sequence peptide epitopes are usually quite easily mapped.[30,31,36,52–54] There are also examples of the identification of peptide ligands for antibodies that recognize conformation-dependent peptide epitopes[55–57] and non-peptide epitopes.[58] Peptide ligands for other proteins that normally bind nonpeptidic compounds have also been found with phage libraries. Examples are peptides that compete with the binding of biotin to streptavidin[32,59] and peptides that displace α-methylmannoside from the lectin, concanavalin A.[60,61] Smith *et al.*[62] recovered from a phage library peptides that are competitive inhibitors of the interaction of S-peptide with S-protein (a fragment of RNase). These peptides have little or no sequence similarity to the natural ligand S-peptide. A family of peptides binding to gpIIb/IIIa was obtained by screening the integrin with a library of random hexapeptides flanked by cysteine residues, capable of forming a disulfide-bridged loop containing the variable hexamer region. The peptide ligands selected contained the tripeptide sequence, RGD, and were much more active in binding the integrin when prepared in the cyclic, oxidized form.[63] Another integrin, $\alpha_5\beta_1$, was screened with a phage hexapeptide library, and most of the sequences recovered contained the RGD motif.[64] One of these, having the potential of forming a disulfide-constrained cyclic structure, was the most potent inhibitor of $\alpha_5\beta_1$ binding to fibronectin. By screening a cell surface antigen receptor derived from the B-cells of a lymphoma patient, Renschler *et al.*[65] identified several families of peptides that compete with an anti-idiotype IgG prepared against the receptor. One of the peptides, when dimerized or tetramerized, activated the receptor leading to the specific effects of tyrosine phosphorylation, apoptosis, and death in culture of cells derived from the tumor. Hammer *et al.*[66] have employed phage peptide libraries to define the consensus binding motifs of three alleles of the MHC class II peptide receptor. Blond-Elguindi *et al.*[67] analyzed the peptide-binding preferences of another promiscuous receptor, the chaperonin BiP. The motif identified allowed the formulation of rules that predicted additional peptide ligand sequences in proteins that interact with BiP.

Phage–peptide libraries have been actively used against receptors for several years now. Because peptides discovered as ligands to pharmaceutically important receptors serve only as very early leads however, very little information has been made publically available concerning the successful identification of novel, low molecular weight peptides for important receptor targets.

Besides their use as sources of ligands, peptide libraries can be adapted, with a little ingenuity, to the selection of peptides with other properties. Matthews and Wells[68] recently described a method for identifying protease substrates using phage libraries. A library of peptides fused to pIII was constructed with the additional feature of a "tether" added at the N-terminus of the peptide-pIII fusion. The tether was chosen to be a high-affinity ligand for some binding protein (these workers used hGH as their tether). The library phage, all of which possessed the tether, were bound to the hGH receptor previously coated onto microtitre wells. The immobilized phage were then treated with the protease subtilisin, and phage-bearing peptides cleaved by the protease were released from their tether and freed from the support. The released phage were recovered for amplification and DNA sequencing. A similar method, utilizing a small peptide epitope as the tether and allowing the protease digestion to occur in solution, has been independently developed and utilized to identify new substrates for a matrix metalloprotease.[69] Substrates for peptide-modifying enzymes other than proteases can be identified with peptide libraries if an appropriate antibody or other binding protein specific for either the substrate or the product of the reaction is available (*vide infra*).

Peptides-on-Plasmids. In each of the filamentous phage systems and the LamB system described above, the physical link between the peptide and its encoding DNA occurs by the containment of the DNA within a particle (cell or phage) that carries the peptide on its surface. Capturing the peptide captures the particle and the DNA within. Cull *et al.*[70] devised an alternative scheme that utilizes the DNA-binding protein LacI to form a link between peptide and DNA. They constructed a plasmid containing the LacI gene with an oligonucleotide cloning site at its 3′-end. Under the controlled induction by arabinose, a LacI–peptide fusion protein is produced. This fusion retains the natural ability of LacI to bind to a short DNA sequence known as the Lac operator (LacO). By installing two copies of LacO on the expression plasmid, the LacI–peptide fusion binds tightly to the plasmid that encoded it. Because the plasmids in each cell contain only a single oligonucleotide sequence and each cell expresses only a single peptide sequence, the peptides become specifically and stably associated with the DNA sequ

that directed its synthesis. The cells of the library are gently lysed and the peptide–DNA complexes are exposed to a matrix of immobilized receptor to recover the complexes containing active peptides. The associated plasmid DNA is then reintroduced into cells for amplification and DNA sequencing to determine the identity of the peptide ligands. As a demonstration of the practical utility of the method, a large random library of dodecapeptides was made and selected on a monoclonal antibody raised against the opioid peptide dynorphin B. A cohort of peptides was recovered, all related by a consensus sequence corresponding to a six-residue portion of dynorphin B.[70]

This scheme, named peptides-on-plasmids by the designers, differs in two important ways from the phage display methods. First, the peptides are attached to the C-terminus of the fusion protein, resulting in the display of the library members as peptides having free carboxy termini. Both of the filamentous phage coat proteins, pIII and pVIII, are anchored to the phage through their C-termini, and the guest peptides are placed into the outward-extending N-terminal domains. In some designs, the phage-displayed peptides are presented right at the amino terminus of the fusion protein.[30] A second difference is the set of biological biases affecting the population of peptides actually present in the libraries. The LacI fusion molecules are confined to the cytoplasm of the host cells. The phage coat fusions are exposed briefly to the cytoplasm during translation but are rapidly secreted through the inner membrane into the periplasmic compartment, remaining anchored in the membrane by their C-terminal hydrophobic domains, with the N-termini, containing the peptides, protruding into the periplasm while awaiting assembly into phage particles. The peptides in the LacI and phage libraries may differ significantly as a result of their exposure to different proteolytic activities. The phage coat proteins require transport across the inner membrane and signal peptidase processing as a prelude to incorporation into phage. Certain peptides exert a deleterious effect on these processes and are underrepresented in the libraries.[71] These particular biases are not a factor in the LacI display system.

Schatz[72] has used the peptides-on-plasmid system to discover a new enzyme substrate: a completely novel and much smaller recognition sequence for the protein-biotinylating enzyme, BirA. Functional homologs of BirA are found in species ranging from bacteria to tomatoes. In *E. coli*, the enzyme adds biotin to the epsilon amino group of a specific lysine in the 75 residue enzyme recognition domain of a single protein made by the bacterium. The recognition domain is highly conserved among prokaryotic and eukaryotic species that possess the BirA-like activity, and attempts to reduce the size of the domain by deletion analysis have not been fruitful. Schatz designed a library with a fixed lysine flanked by 10 random residues on each side and fused to the C-terminus of LacI. The LacI fusions are expressed in the bacterial cytoplasm where the peptides are exposed to the BirA enzyme. The LacI–plasmid complexes were isolated in the usual way and the library screened for binding to immobilized streptavidin to retain peptides that had become biotinylated. An important counter selection was employed to bias against the recovery of non-biotinylated peptides that contained streptavidin-binding motifs of the types reported by Devlin[32] and Kay.[59] The counter selection was based on the assumption that the nonbiotinylated motifs would bind streptavidin with much lower affinity than the peptides containing a biotinylated lysine. After binding the plasmid library to immobilized streptavidin, biotin was added to displace the low-affinity streptavidin-binding peptides. Sequences of the clones that survived this process revealed a 13-residue consensus around the recipient lysine that served as a substrate for the BirA enzyme.

Peptide-modifying enzymes can be used to create additional diversity in the libraries. To illustrate, libraries of C-terminally amidated peptides have been constructed. The dual activity enzyme, peptidyl amino monooxygenase (PAM), cleaves the C-terminal glycine of peptides to leave a carboxamide on the penultimate residue. Because the enzyme is quite permissive in its specificity for the residues adjacent to the glycine, most peptides should serve as substrates for this transformation. A library was designed to contain 12 random residues followed by a single fixed glycine, all fused to the C-terminus of LacI. The LacI–peptide–plasmid complexes were isolated and treated with PAM *in vitro*, providing display of a collection of random peptides terminating with an amide rather than a free carboxylic acid residue.[73] This type of library may be a fruitful source of new ligands for the family of G-protein receptors that have C-terminally amidated peptides as their natural ligands.

Peptides Displayed on Polysomes: An *in Vitro* Approach to Random Library Construction. The number of small peptides available in recombinant random libraries is enormous. Libraries of 10^7–10^9 independent clones are routinely prepared. Libraries as large as 10^{11} recombinants have been created,[74] but this size approaches the practical limit for clone libraries. This limitation in library size occurs at the step of transforming the DNA containing randomized segments into the host bacterial cells. To circumvent this limitation, an *in vitro* system based on the display of nascent peptides in polysome complexes has recently been developed. This new addition to the repertoire of L-amino acid display libraries has the potential of producing libraries 3–6 orders of magnitude larger than the currently available phage/phagemid or plasmid libraries. The salient feature of this approach is that the construction of the libraries, expression of the peptides, and screening, is done in an entirely cell-free format.

As a demonstration of this method, Mattheakis *et al.*[75] have constructed a molecular DNA library encoding 10^{12} decapeptides and have expressed the library in an *E. coli* S30 *in vitro* coupled transcription/translation system. Conditions were chosen to stall the ribosomes on the mRNA, causing the accumulation of a substantial proportion of the RNA in polysomes and yielding complexes containing nascent peptides still linked to their encoding RNA. The polysomes are sufficiently robust to be affinity purified on immobilized receptors in much the same way as the more conventional recombinant peptide display libraries are screened. RNA from the bound complexes is recovered, converted to cDNA, and amplified by PCR to produce a template for the next round of synthesis and screening. The polysome display method can be coupled to the phage display system, as shown in this report. Following several rounds of screening, cDNA from the enriched pool of polysomes was cloned into a phagemid vector. This vector serves as both a peptide expression vector, displaying peptides fused to the coat proteins, and

as a DNA sequencing vector for peptide identification. By expressing the polysome-derived peptides on phage, one can either continue the affinity selection procedure in this format or assay the peptides on individual clones for binding activity in a phage ELISA (*vide infra*), or for binding specificity in a competition phage ELISA.[76] To identify the sequences of the active peptides one simply sequences the DNA produced by the phagemid host.

The authors screened the polysome library for ligands to the anti-dynorphin B antibody, 32.39. This antibody recognizes as its epitope a six-residue portion of dynorphin B, RQFKVV, and has been shown to bind a large number of peptides closely related to this sequence.[70] Three to five rounds of screening yielded a family of peptides clearly related to the known ligands and having affinities, as free peptides, of 7–140 nM. The selection of peptides in this affinity range indicates the polysomes may form a monovalent association with the immobilized receptors (see discussion below on the effect of valency in library selection schemes). The molecular library described, 10^{12} DNA molecules, was expressed in an *in vitro* transcription/translation reaction of 50 μL. Reactions 100-fold larger are easily attained, and modifications to the S30 system could also provide additional synthetic capacity. It may therefore be possible to create and screen polysome libraries of 10^{14}–10^{15} members. The peptide population of the polysome libraries should be less influenced by the biological biases discussed above, and additional building block diversity might potentially be employed by incorporating tRNAs charged with unnatural amino acids into the reaction, as pioneered by Schultz *et al.*[77]

Designing Random Libraries for Biological Display Systems. Frequently, a drug-discovery program may be initiated with no information about the structural requirements of a small ligand for a target receptor. This is especially true for programs in which the goal is to disrupt protein/protein interactions. The large families of cytokine and growth factor receptors are examples of targets of this type. Historically, there have been few examples of low molecular weight compounds that block such interactions. For receptors where small molecule leads are available, it may still be desirable to identify new leads that bind to the target in a manner that is either similar or dissimilar to that of a known ligand. A likely strategy for finding such ligands may be searching the huge collections of compounds available in randomly created peptide display libraries. In this section the salient points in the design of these libraries are discussed.

In the construction of synthetic libraries, the choice of appropriate building blocks from among the huge number of available monomers can potentially be problematic. This is usually not an issue in the construction of random biological peptide libraries, where one simply uses all of the 20 genetically encoded amino acids. More pressing design issues relate to the length of the peptides to be displayed and the fixing of amino acids in certain positions—often for the purposes of introducing structural constraint into the peptides. Choosing a peptide length so short that the library size greatly exceeds the number of peptides theoretically accessible is obviously a defective strategy. If, for example, a library of L-tetrapeptides were to be made (1.6×10^5 possible tetrapeptides, $\sim 10^6$ possible tetracodons), there would be little point in constructing the library to contain 10^8 members. Often the goal in library design is to create the greatest number of *different* compounds. This requires making a library of the largest practical size and choosing a peptide length whose combinatorial possibilities greatly exceed the library size. This results in a low likelihood that the same peptide will be made by more than one clone of the library. To illustrate, there are 2.6×10^{10} L-octapeptides; even the more abundant peptides comprised entirely of 3-codon amino acids are unlikely to appear more than once in a library of 10^8 members; thus, this library may express nearly 10^8 different peptides. Of course, peptides containing many of the 12 amino acids encoded by a single codon will be poorly represented in this library. To include most of the octapeptides, a library exceeding 10^{12} members is required.

The nature of the target receptor should also be considered when designing the library. Although some binding sites for small ligands are known to bind peptides of two or three residues, other important targets, especially those involved in protein/protein interactions, may prefer larger molecules as ligands. Libraries containing peptides of 8, 10, 12 or more random residues may be the most productive for early screening of receptors of this type.

Biological display libraries may be designed to contain randomized residues placed within the context of fixed "scaffolds" of constraining residues. Obvious examples include libraries of disulfide-bridged loops, made by flanking strings of random residues with pairs of fixed cysteines. Restricting the global conformations of peptides may provide ligands of higher affinity that are more amenable to medicinal chemical development, but this may come at the cost of *effective* library size. Each member of a peptide library exists as a family of conformational states, and as one deliberately constrains the peptides, the universe of molecular shapes to be sampled by the library declines. To exploit the conformational diversity in unconstrained libraries requires screening techniques that, in the early rounds, can recover rather low affinity ligands. Some of the multivalent, high-avidity systems described above are capable of selecting ligands having K_d values as weak as 100 μM. Early, low-affinity leads can then be improved by using analog libraries and affinity-sensitive screening strategies as described in later sections of this review.

Designing Libraries of Variants Based on a Lead Peptide. Often, one has some knowledge about the primary structures of small peptide ligands of the target receptor. This information may come from the structure of the receptor, from the structure of the natural ligand, or from early leads identified in random library screening. This information can be used to construct libraries of analogs ("mutants") based on the starting sequence ("wild type"). Screening the analog libraries aids in determining the minimum size of the active sequence and in identifying those residues critical for binding and intolerant of substitution. The mutant libraries may also contain peptides of higher affinities than the starting sequence, and these are recovered with specialized screening techniques (*vide infra*).

Several methods are available to construct analog libraries. Beginning with a single peptide sequence, or with a consensus sequence representing a group of related peptides, the encoding oligonucleotide is resynthesized, and a controlled level of incorrect bases is introduced at each coupling step.[78–80] Both low and high levels of mutagenesis are feasible. A typical protocol calls for a

mixture of 70% of the correct base and 10% of each of the other three bases (70/10/10/10 strategy) and produces amino acid changes in about 50% of the positions. The resulting mixture of oligonucleotides is cloned into the display vector to create a library of 10^8–10^9 variants for rescreening. The method is very convenient but somewhat biased in the distribution of changes produced from the starting sequence. For example, a Val to Ala change requires that only one base of the codon change (GTN to GCN), an event with 0.1 probability. By contrast, a Asp to Trp change requires a different base in all three codon positions (GAT/C to TGG), an event with only 0.001 likelihood. Nevertheless, because very large libraries can be made and very rare events detected, this approach is highly productive in defining the critical residues involved in binding.

A less biased, but technically more difficult alternative is the codon-by-codon mutagenesis strategy. The pool of synthesis resin is apportioned between two (or more) reactions, one dedicated to making the wild-type codons, and the other(s) to making random codons (or other specific amino acid codons). The ratio of wild-type to mutant codon is determined by the proportion of the resin placed in each reaction. At each codon boundary, the resin is pooled, mixed, and reapportioned to the reaction chambers.

Both of these methods cause mutations to be distributed throughout the starting peptide at a controllable frequency. A different approach fixes certain residues and randomizes the others. One might make a mutant library based on an octapeptide lead by fixing the first four residues as those in the starting sequence and randomizing positions 5–8. A complementary library fixing positions 5–8 and randomizing the first four residues would also be made. From screening these libraries, much can be learned about the preferences of the receptor for each position in the ligand sequence. Of course one could make all 70 libraries of 8 residues taken 4 ways ($n!/[x!(n-x)!]$), but enough information is usually obtained with a very few such libraries. A variation on this theme may be employed when, in early screening, an extraordinary conservation of particular amino acids at certain positions is observed. A sublibrary is made to fix (or lightly mutate) these residues and randomize the remaining positions.

Finally, short, random regions can be added to the starting sequence. Although, in general, reducing the size of the initial lead is desirable, peptides of higher affinity may be obtained from libraries constructed to contain short random regions flanking the starting sequence. Selection of the higher affinity members of mutagenesis libraries requires the application of screening conditions sensitive to intrinsic affinity differences among the library members, as discussed in Part 2 of this Perspective (scheduled to appear in *J. Med. Chem.* 1994, *37* (10)). The construction and screening of mutagenesis libraries serves to compile structure–activity relationships around analogs of a starting compound. As such, the process is smoothly coupled to the parallel synthetic library techniques described in the next section, extending the array of building block monomers utilized in the analog constructions.

C. Synthetic Chemical Approaches to Generating Peptide Diversity

During the last two decades peptide chemistry has been steadily attaining ever greater prominence in pharmaceutical research, fueled, in part, by the discovery of many biologically active peptide hormones and an increased understanding of the roles that these molecules play in regulating a multitude of human physiological responses. Consequently many low molecular weight bioactive peptides have become the focus of medicinal chemistry research efforts to develop hormone receptor agonists or antagonists.[1] In addition, peptidic structures commonly are found in molecules designed to inhibit enzymes that catalyze proteolysis, phosphorylation, and other post-translation protein modifications that may play central roles in the pathologies of various disease states.[81]

The desire to develop and explore structure–activity relationships around peptide lead compounds has placed tremendous demands on the productivity of peptide chemistry. Over the last 10 years a variety of methodologies have been developed that permit simultaneous synthesis of multiple peptides, and this area has recently been thoroughly reviewed.[82] Because many of the technologies underlying the construction of diverse synthetic peptide libraries are rooted in multiple peptide synthesis (MPS) strategies, a brief overview of the principal MPS methods and their applications is provided below.

Multiple Peptide Synthesis. The various approaches to simultaneous preparation and analysis of large numbers of synthetic peptides each rely on the fundamental concept of synthesis on a solid support introduced by Merrifield in 1963.[83] Generally, these techniques are not dependent on the protecting group or activation chemistry employed, although most workers today avoid Merrifield's original ^{t}Boc/Bzl strategy in favor of the more mild Fmoc/^{t}Bu chemistry and efficient hydroxybenzotriazole-based coupling agents. Many types of solid matrices have been successfully used in MPS, and yields of individual peptides synthesized vary widely with the technique adopted (e.g., nanomoles to millimoles). Interestingly, the early architects of these methodologies were molecular immunologists, primarily concerned with determining the immunodominant B-cell epitopes of viral proteins and in deducing the fine specificity of peptide recognition by monoclonal antibodies.

(A) Multipin Synthesis. Geysen and co-workers[84] introduced a method for generating 96 peptide sequences by a parallel synthesis on polyacrylic acid-grafted polyethylene pins arrayed in the microtitre plate format. In the original experiments ~50 nmol of (ideally) a single peptide sequence was covalently linked to the spherical head of each pin, and interactions of each peptide with monoclonal antibodies or polyclonal sera could be determined in a direct binding assay by ELISA. The Geysen group claims to be able to synthesize and screen thousands of peptides per week using the multipin method, and the tethered peptides may be reused in many assays. In subsequent work the level of peptide loading on individual pins was increased to as much as 2 μmol/pin by grafting greater amounts of functionalized acrylate derivatives to detachable pin heads.[85] Appropriate linker moieties have also been appended to the pins so that the peptides may be cleaved from the supports after synthesis for assessment of purity and evaluation in competition binding or functional bioassays.[86–88]

The most significant applications of the multipin technology have been in the area of anti-protein antibody epitope analysis[89,90]. Continuous epitopes are readily mapped by synthesizing on pins a complete set of

overlapping peptides derived from the antigen's primary structure and measuring the reactivity of the antibody or antiserum by ELISA (Pepscan technique[89]). For example, a protein antigen of 200 amino acids may be scanned by evaluating antibody interactions with 193 overlapping octapeptide fragments. Upon its identification, the fine specificity of an antibody epitope can be further deduced by "replacement set analysis" in which the effect on antibody recognition of substituting each of the amino acids within the epitope by other amino acids is observed. Those residues which poorly tolerate substitution are implicated as being directly involved in antigen–antibody interaction. Antibody binding sites on a wide range of viral proteins have been elucidated by the Pepscan method.[82,84,91-93] More recent applications of the multipin method of MPS have taken advantage of the cleavable linker strategy to prepare soluble peptides for T-cell proliferation studies.[94-96] An investigation of the binding activity of 512 stereoisomers of the tachykinin peptide, substance P, generated by synthesizing all possible D-amino acid replacements at 9 of the 11 positions in this neuropeptide was recently reported.[97,98] Radioligand binding experiments indicated that while substitutions toward the C-terminus of the peptide generally abrogated NK1 receptor binding, substitutions in the N-terminal region were often well tolerated. In structure–activity relationship studies of a hexapeptide endothelin receptor antagonist conducted by Spellmeyer *et al.*, multipin synthesis was used in the systematic substitution of 50 different amino acids at each position in the lead peptide.[99,100] Competition binding assays run with crude peptides cleaved from the pins provided an initial screen from which a number of more potent analogs were identified.

(B) Tea-Bag Method. In the so-called "tea-bag" MPS method first developed by Houghten, peptide synthesis occurs on resin that is sealed inside porous polypropylene bags.[101] Amino acids are coupled to the resins by placing the bags in solutions of the appropriate individual activated monomers, while all common steps such as resin washing and α-amino group deprotection are performed simultaneously in one reaction vessel. At the end of the synthesis, each bag contains a single peptide sequence, and the peptides may be liberated from the resins using a multiple cleavage apparatus.[102] This technique offers advantages of considerable synthetic flexibility and has been partially automated to permit the simultaneous synthesis of up to 150 different peptides.[103] Moreover, soluble peptides of greater than 15 amino acids in length can be produced in sufficient quantities (*ca.* 500 μmol) for purification and complete characterization if desired.

Multiple peptide synthesis using the tea-bag approach has been applied to a range of molecular recognition problems including antibody epitope analysis,[101] peptide hormone structure–function studies,[104,105] and protein conformational mapping.[106] For example, using a combination of alanine and D-amino acid scanning, the C-terminal tetrapeptide fragment of the 36-amino acid peptide hormone, neuropeptide Y, has been shown to be essential for receptor recognition. Neuropeptide Y analogs that link regions of the amino and carboxy termini of the parent peptide via a spacer residue were shown to behave as receptor agonists, mimicking a discontinuous pharmacophore of the hormone.[104,105] Families of synthetic peptides modeled around the Arg-Gly-Asp and γ chain fragments of fibrinogen have been prepared using the tea-bag method to define antagonists of the platelet integrin receptor gpIIb/IIIa.[107]

(C) Multiple Peptide Synthesis through Coupling of Amino Acid Mixtures. Simultaneous coupling of mixtures of activated amino acids to a single resin support has been used as a multiple peptide synthesis strategy on several occasions.[80,108-111] For example, four to seven analogs of the magainin 2 and angiotensinogen peptides were successfully synthesized and resolved in one HPLC purification after coupling a mixture of amino acids at a single position in each sequence.[108] This approach has also been used by two groups to prepare degenerate peptide mixtures for defining the substrate specificity of endoproteolytic enzymes.[110,111] In these experiments a series of amino acids was substituted at a single P′ position within the substrate sequence. After proteolysis, Edman degradation was used to quantitate the yield of each amino acid component in the hydrolysis product and hence to evaluate the relative k_{cat}/K_m values for each substrate in the mixture.

The operational simplicity of synthesizing many peptides by coupling monomer mixtures is offset by the difficulty in controlling the composition of the products. The product distribution reflects the individual rate constants for the competing coupling reactions, with activated derivatives of sterically hindered residues such as valine or isoleucine adding at a significantly slower rate than glycine or alanine for example. The nature of the resin-bound component of the acylation reaction also influences the addition rate, and the relative rate constants for the formation of 400 dipeptides from the 20 genetically coded amino acids have been determined by Rutter and Santi.[109] These reaction rates can be used to guide the selection of appropriate relative concentrations of amino acids in the mixture to favor more closely equimolar coupling yields.

(D) Multiple Peptide Synthesis on Nontraditional Solid Supports. The search for innovative methods of multiple peptide synthesis has led to the investigation of alternative polymeric supports to the polystyrene–divinylbenzene matrix originally popularized by Merrifield. Cellulose, either in the form of paper disks[112-115] or cotton fragments[116,117] has been successfully functionalized for peptide synthesis. Typical loadings attained with cellulose paper range from 1 to 3 μmol/cm^2, and HPLC analysis of material cleaved from these supports indicates a reasonable quality for the synthesized peptides. Alternatively, peptides may be synthesized on cellulose sheets via noncleavable linkers and then used in ELISA-based receptor binding studies.[118] The porous, polar nature of this support may help suppress unwanted nonspecific protein binding effects. By controlling the volume of activated amino acids and other reagents spotted on the paper, the number of peptides synthesized at discreet locations on the support can be readily varied. In one convenient configuration spots are made in an 8 × 12 microtiter plate format. Frank has used this technique to map the dominant epitopes of an antiserum raised against a human cytomegalovirus protein, following the overlapping peptide screening (Pepscan) strategy of Geysen.[118] Other membrane-like supports that have been investigated for multiple solid-phase synthesis include polystyrene-grafted polyethylene films.[119]

Automated Multiple Peptide Synthesis Methods. There has been considerable effort devoted to constructing

automated multiple peptide synthesizers, and several designs have now been successfully commercialized. The groups of Schnorrenberg[120] and Gausepohl[121] have employed pipetting robots to deliver solvents, activated amino acids and other reagents to peptide resins contained within an array of 96 and 48 reaction vessels, respectively. Individual peptides may be synthesized on a scale of 10–50 μmol using these instruments. Zuckermann *et al.* have designed a related robotic delivery system that can simultaneously prepare 36 peptides (at up to 125 μmol per peptide) when operated in multiple synthesis mode.[122] This workstation has the additional capability of mixing and redistributing resin from multiple reaction vessels which significantly facilitates the assembly of combinatorial peptide libraries (*vide infra*). Although the operation of these instruments differs in details such as resin agitation and washing, similar synthesis chemistries can be utilized with each, and the purities and yields of the resulting peptides obtained are typically respectable.

Synthesizers based on the continuous flow principle[123] have also been successfully used to prepare milligram quantities of multiple peptides.[113,124,125] Novel, porous polymer supports including cellulose and poly(dimethylacrylamide) have typically been used in these column reactor formats. Other innovations in multiple synthesis relate to strategies for automating the deprotection and cleavage of many peptides in parallel.[126,127] An integrated multichannel synthesizer and cleavage device has recently been described by Neimark and Briand.[127]

Combinatorial Synthetic Peptide Libraries. The generation of molecular diversity using strategies that covalently connect together members of a set of chemical building blocks in all possible combinations represents a revolution in multiple synthesis. Parallel development of a variety of new high-throughput screening methodologies has made evaluation of these combinatorial libraries in biological assays both practical and efficient. The various technologies that have emerged for generating and screening peptide libraries may be fundamentally distinguished by the format in which the diversity is presented (tethered vs soluble libraries; physically segregated ligands vs mixed pools). This in turn dictates the type of biological assay methods that may be utilized, influences the strategy followed in ligand structure elucidation, and ultimately determines the size of libraries that may be practically screened. The common hallmark of these combinatorial techniques is that within each chemical coupling step, multiple compounds are generated simultaneously such that each synthesis cycle results in an exponential increase in library size.

Two distinct mechanisms for elaborating molecular diversity may be defined. In the first, mixtures of activated monomers are coupled to one (or more) solid supports at each cycle of the synthesis.[90,109,128,129] As previously noted, the product distribution here is influenced by the relative kinetics of the competing reactions, making it difficult, in general, to ensure an equimolar representation of each component in the library. Ligand equimolarity is an important criterion if one wishes to make even semiquantitative determinations of the affinities of receptor–ligand interactions by equilibrium measurements with compound mixtures. The second mechanism circumvents the problem of competing coupling reactions by physically segregating the support into multiple aliquots or spatially discreet fractions to which addition of a single monomer

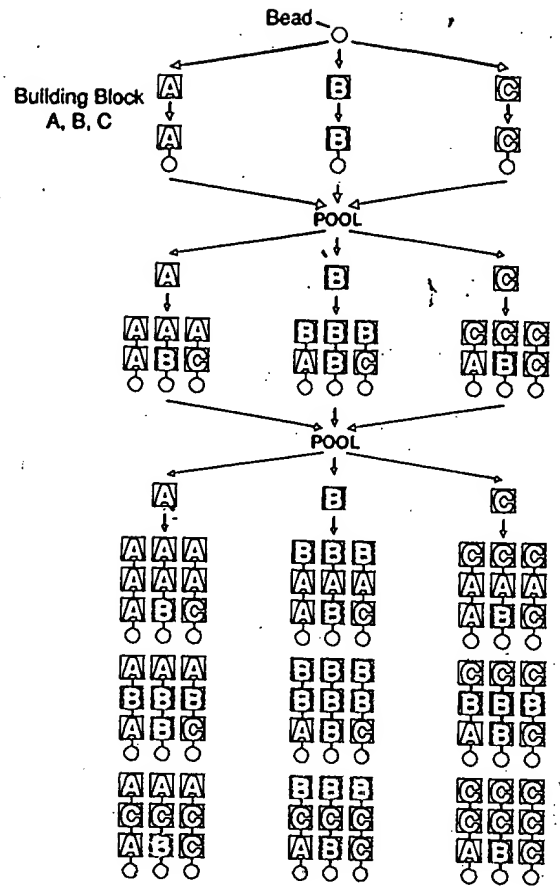


Figure 2. Preparation of a combinatorial library by the split synthesis method.

can occur. The earliest description of this approach was the "portioning–mixing" method of peptide synthesis of Furka and co-workers,[130–132] and termed the "split synthesis"[133] or "divide, couple, and recombine" process[134] by other workers. This simple procedure involves dividing the resin support into n equal fractions, coupling a single monomer to each aliquot in a separate reaction, and then thoroughly mixing all the resin particles together. Repeating this protocol for a total of x cycles can produce a stochastic collection of up to n^x different peptides, as governed by the Poisson distribution. A schematic illustration of a split synthesis using three building blocks A, B, and C to generate all 27 possible trimer combinations is shown in Figure 2.

Combinatorial Libraries Using Multipin Synthesis. The problem of finding peptide ligands that bind to monoclonal antibodies which recognize discontinuous epitopes within a protein antigen was first approached by Geysen using a combinatorial synthesis procedure referred to as the "mimotope strategy".[89,90] Peptide mixtures are synthesized on pins in the format *-*-*-D_1-D_2-*-*-* where the positions D_1 and D_2 are occupied by single defined amino acids and the * positions represent randomly incorporated residues resulting from coupling mixtures of activated amino acids. Using the 20 common L-amino acid monomers, a total of 400 such mixtures may be synthesized, each ideally containing $20^6 = 6.4 \times 10^7$ different sequences. The peptide mixture on each pin is screened for antibody binding by ELISA to identify an optimum dipeptide sequence X-Y. This sequence then provides the basis for a further round of synthesis in the format *-*-D-X-Y-*-*-*, where the number of degenerate positions is reduced by one, allowing resolution of a third

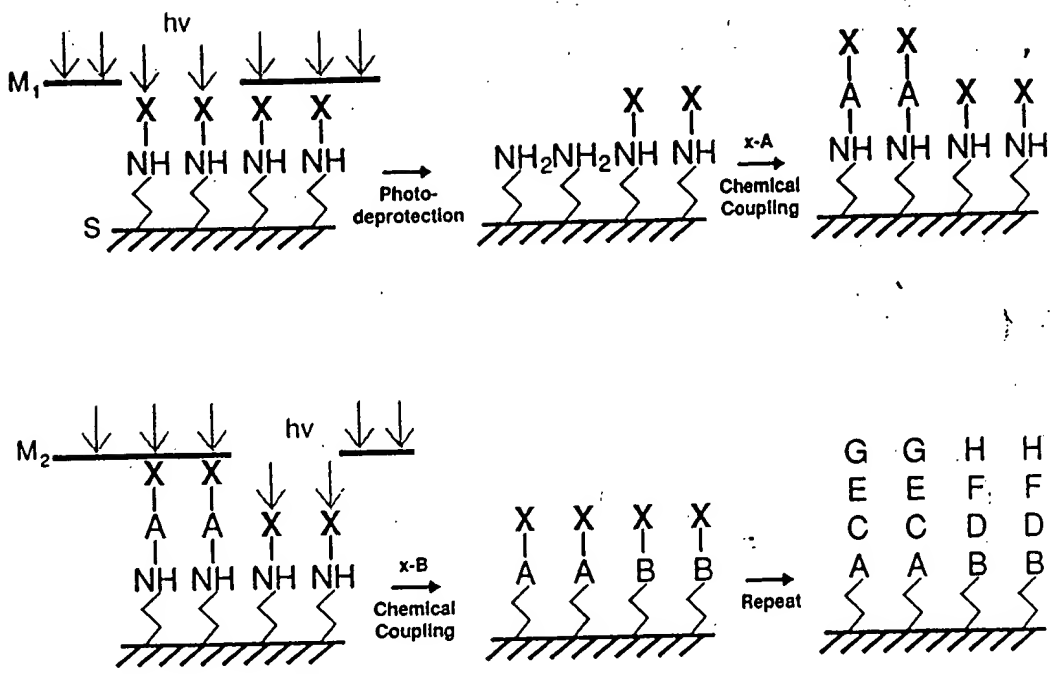


Figure 3. Concept of light-directed spatially addressable parallel chemical synthesis. A substrate S bears amino groups that are blocked with a photolabile protecting group X. Illumination of specific regions through a lithographic mask M_1 leads to photodeprotection. Amino groups in the exposed sector of the substrate are now accessible for coupling (for example by the Merrifield solid-phase peptide synthesis method). The first chemical building block A containing a photolabile protecting group X is then attached. A different mask M_2 is used to photoactivate a different region of the substrate. A second labeled group X–B is added and condensed to the newly exposed amino groups. Additional cycles of photodeprotection and coupling are carried out to obtain the desired set of products.

residue in the sequence. This cycle of synthesis and screening is reiterated until the entire sequence is optimized, and the resulting peptide is termed a "mimotope" (for epitope mimetic) for the antibody.

Geysen has recently reviewed use of the mimotope strategy for the identification and optimization of ligands for antibodies and other receptors.[135] A preferred approach is to screen hexapeptides that incorporate both L- and D-amino acids ($39^6 = 3.52 \times 10^9$ hexamers from the common monomers), though non-α or other nonnatural amino acids can certainly be used in these syntheses.

Combinatorial Libraries by Light-Directed, Spatially Addressable Parallel Chemical Synthesis. A scheme of combinatorial synthesis in which the identity of a compound is given by its location on a synthesis substrate is termed a spatially addressable synthesis. Here, the combinatorial process is carried out by controlling the addition of a chemical reagent to specific locations on a solid support.[4,136–138] The technique combines two well-developed technologies: solid-phase peptide synthesis chemistry and photolithography. The high coupling yields of Merrifield chemistry allow efficient peptide synthesis, and the spatial resolution of photolithography affords miniaturization. The merging of these two technologies is done through the use of photolabile amino protecting groups in the Merrifield synthetic procedure.

The key points of this technology are illustrated in Figure 3. A synthesis substrate is prepared for amino acid coupling through the covalent attachment of photolabile nitroveratryloxycarbonyl (NVOC) protected amino linkers. Light is used to selectively activate a specified region of the synthesis support for coupling. Removal of the photolabile protecting groups by light (deprotection) results in activation of selected areas. After activation, the first of a set of amino acids, each bearing a photolabile protecting group on the amino terminus, is exposed to the entire surface. Amino acid coupling only occurs in regions that were addressed by light in the preceding step. The solution of amino acid is removed, and the substrate is again illuminated through a second mask, activating a different region for reaction with a second protected building block. The pattern of masks and the sequence of reactants define the products and their locations. Since this process utilizes photolithographic techniques, the number of compounds that can be synthesized is limited only by the number of synthesis sites that can be addressed with appropriate resolution. The position of each compound is precisely known; hence, its interactions with other molecules can be directly assessed. The assay molecule can be labeled with a fluorescent reporter group to facilitate the identification of specific interactions with individual members of the matrix. This technology requires only micropreparation and consumption of chemical reagents as well as micropreparation and consumption of biological targets.

The light-directed synthesis of an array of two peptides YGGFL and PGGFL was performed. Following synthesis, the pentapeptide array was probed with a mouse monoclonal directed against β-endorphin (mAb 3E7). This antibody requires an amino-terminal tyrosine residue for high affinity recognition. A second incubation with fluorescein-labeled goat-anti-mouse IgG was used to label the regions containing bound 3E7. The array was then scanned in an epifluorescence microscope. The result of the scan is shown in Figure 4. A high-contrast fluorescence checkerboard image shows that YGGFL and PGGFL were synthesized in alternating 50-μm squares, that the YGGFL synthesized on the surface is accessible for binding to the antibody 3E7, and that the antibody does not recognize PGGFL. The contrast between synthesis sites is very high at 50-μm resolution. This spatial resolution provides 40 000 discrete synthesis sites in a 1-cm × 1-cm square.

In a light-directed chemical synthesis, the products depend on the pattern of illumination and on the order

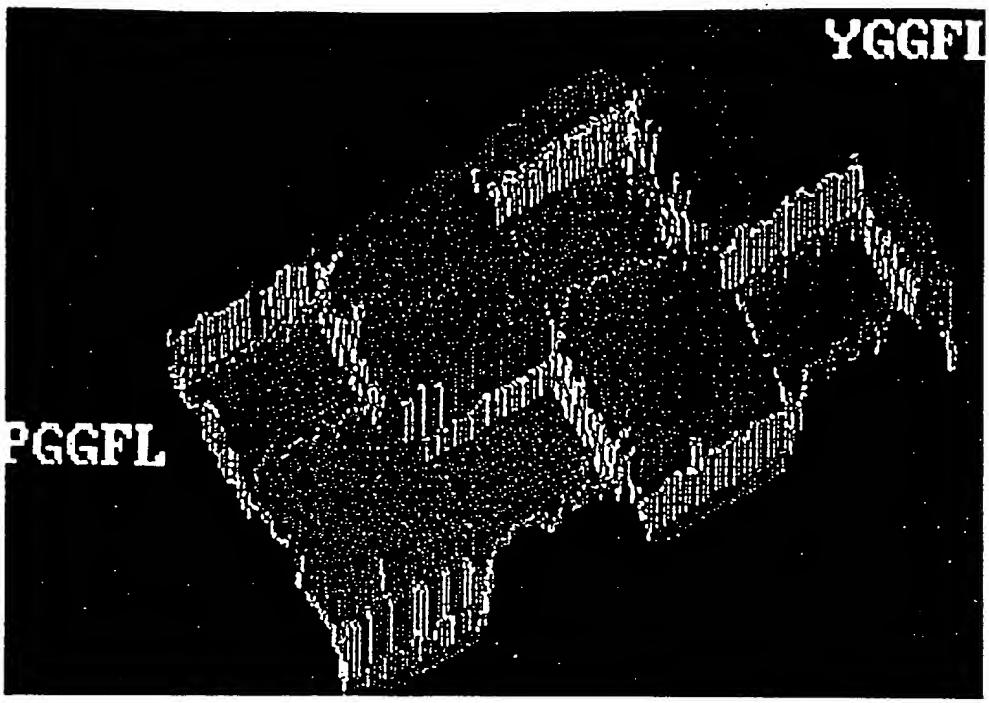


Figure 4. Detection of binding events with laser confocal fluorescence microscopy. After synthesis, the array of compounds is mounted on a thermostatically regulated flow cell and then interrogated with fluorescently-labeled target molecules. Binding events are detected by their fluorescence emissions, which are scored with a photomultiplier tube as described in the text.

of addition of reactants. By varying the lithographic patterns, many different sets of compounds can be synthesized in the same number of steps; this leads to the generation of many different masking strategies. As an example, an eight-step synthesis is shown in Figure 5. The reactants are the ordered set {A, B, C, D, E, F, G, H}. In the first cycle, illumination through photolithographic mask M_1 activates the left one-quarter of the synthesis area. Building block A is added and reacts only at the previously illuminated region. Illumination through M_2 activates the second one-quarter of the substrate, followed by addition of building block B. This process is continued through the first four cycles and results in the pattern of compounds shown in the first half of Figure 5a. M_5 then illuminates *across* the patterns generated in the first four cycles. This one illumination through M_5 now generates four distinct dimers. The process is continued to generate the set of 16 compounds shown in the Figure 5. The final product set is {AE, AF, AG, AH, BE, BF, BG, BH, CE, CF, CG, CH, DE, DF, DG, DH}.

This strategy can be applied to a large set of building blocks. For example, the complete set of 400 dipeptides can be formed from the set of 20 L- or D-amino acids. Analogous to Figure 5a, the synthesis may be performed via two rounds of 20 coupling cycles per round. In cycle 1 of the first round, M_1 activates 1/20th of the substrate for coupling with the first of 20 amino acids. An additional 19 translations and illuminations followed by 20 coupling cycles are required to complete round 1. The substrate now consists of 20 rectangular stripes each bearing a distinct member of the 20 amino acid "building blocks". The masks of round 2 are perpendicular to round 1 masks, and therefore, a single illumination–coupling cycle in round 2 yields 20 dipeptides. The other 19 cycles in round 2 complete the synthesis of 400 dipeptides.

A general formalism describes the combinatorial synthesis for any spatially addressable chemical synthesis. The process is coveniently expressed in matrix notation[136]

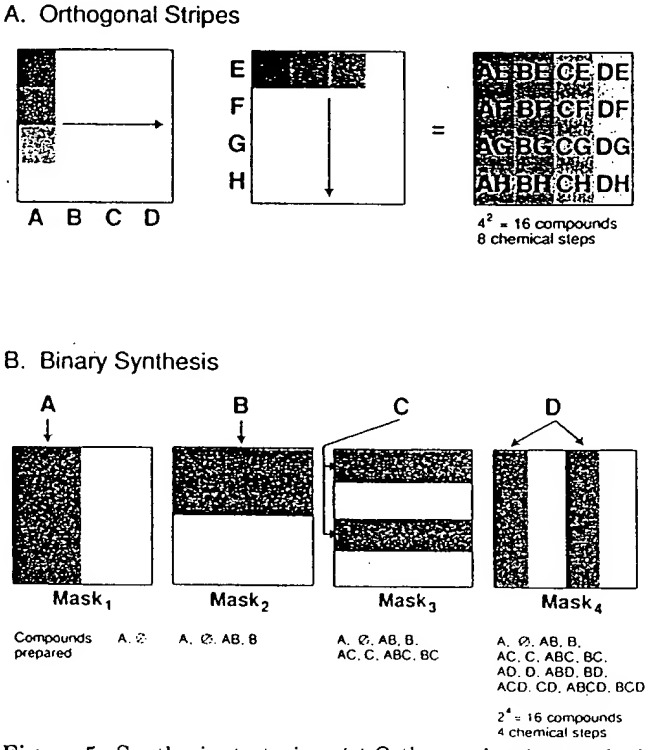


Figure 5. Synthesis strategies. (a) Orthogonal stripe method. Using a strategy similar to the split-resin method, a layer of monomers is formed by photolyzing stripes for each building block. Dimers are formed by photolyzing stripes orthogonal to the first set, preparing n^2 compounds in $2n$ chemical steps. (b) Binary synthesis. Half of the synthesis surface is photolyzed during each coupling step, with subsequent photochemical steps overlapping one-half of the previous synthesis space. With this strategy, 2^n compounds are made in n chemical steps.

and is based on the notion that at any given synthesis site the decision as to whether or not to add a monomer of the building block set is a binary process. If a particular position is to receive a new monomer group, it is addressed

by a 1 (light on condition), if not, it receives a 0. For example, to form ACD from the ordered set {A, B, C, D}, the light switch consists of a vector (switch vector) [1,0,1,1]. Each product element is, therefore, given by the dot product of two vectors, a chemical reactant vector and a binary vector ([A,B,C,D]·[1,0,1,1] = ACD in the previous example). This formalism is very powerful in defining a complex chemical synthesis. For example, consider a four-step synthesis with the ordered set of reactants {A, B, C, D}. Any switch vector consists of four bits for the four chemical reactants. There are 16 four-bit vectors [0,0,0,0] to [1,1,1,1]; hence, a maximum of 16 different products can be made from the ordered set of four reagents (one of the four compounds is the null or zero compound). For the example shown in Figure 5a, there was an ordered set of eight chemical reactants, and 16 compounds were formed. A more efficient synthesis strategy would create all 2^8 (256) possible compounds. Thus, a "binary masking strategy" where the maximum number of compounds are generated in the fewest number of steps when using pure monomers in each step (Figure 5b) was developed. For n number of steps, 2^n number of compounds are generated. Since one of these compounds is the linker where no polymer has been assembled, $2^n - 1$ polymers are made in this strategy. A 16-step binary synthesis, for example, would result in 2^{16} (65 536) regions for screening, of which 65 535 would represent peptides or oligonucleotides. This strategy also creates a distribution of chain lengths, ranging from 1 to n, with the maximum number of polymers being $n/2$ in length. The binary strategy generates all possible permutations of the building blocks while preserving their order of introduction. Hence all deletion sequences possible in the starting sequence are obtained, as well as all possible truncation sequences. This type of strategy has proven useful in epitope mapping experiments with an anti-dynorphin B antibody, D32.39. A binary strategy was employed to generate all permutations about the C-terminus region of dynorphin B, from which RQFKVVT was identified as the minimal binding subunit.[139-141]

There are circumstances where biopolymers in the array should be the same length, and other masking strategies have been developed to accomodate these applications. It may also be desirable to generate sets of biopolymers where all members of a certain class are represented in one array and to rank order the *relative* affinity of each member of the set. A common strategy in ligand optimization is the substitution of a set of novel monomers (e.g., 20 L-amino acids or D-amino acids) for each monomer in a lead compound. This can be accomplished with relatively few steps in a combinatorial synthesis. The number of steps required to substitute x monomers into a polymer of length n is given by $2(n - 1) + x$. For example, substitution of 20 novel amino acids into each position of an octapeptide represents 160 peptides (8×20), which would require up to 1280 individual steps (160×8) if carried out conventionally; combinatorial strategies can create an array of these 160 peptides in 34 reactions.

The maximum number of compounds that can be made in the array is governed largely by the lithographic performance of the system. Practical limitations on spatial resolution, or the desire for a specific class of compounds, may require one to select a particular subset of the possible compounds for a given synthesis. Among other things, the masking resolution is dependent on the wavelength of the light employed and the distance between the mask and the substrate. It has been possible to generate arrays of up to 65 536 compounds in an area slightly greater than 1 cm^2. Further improvements in resolution will allow even higher density arrays to be generated.

Combinatorial Libraries of Soluble Peptides. The "split synthesis" algorithim is readily adapted to generating equimolar mixtures of soluble peptides that may be screened in a variety of competition binding or functional bioassays. This approach has been most extensively investigated by Houghten *et al.*, and bioactive peptides have been identified from libraries containing more than 50 million different sequences.[134,142-146] These hexamer libraries consist of sets of sublibraries that are conveniently prepared by the tea-bag technique in one of two formats: (i) the dual-fixed position series D_1-D_2-*-*-*-* in which a mixture of amino acids is randomly incorporated at four positions of the peptide while two residues (D_i) are explicitly defined, and (ii) the positional scanning series D_1-*-*-*-*-*, *-D_2-*-*-*-*, *-*-D_3-*-*-*, *-*-*-D_4-*-*, *-*-*-*-D_5-*, and *-*-*-*-*-D_6 where single positions are defined throughout the sequence. For libraries constructed from 20 amino acids, 400 distinct dual position sublibraries and 120 positional scanning sublibraries must be synthesized. Active peptides are identified from dual-position libraries by an iterative process of screening and sublibrary resynthesis in a manner that is completely analogous to the mimotope resolution strategy. Thus at each stage of the screening, an additional residue is defined explicitly by evaluating 20 sublibraries in the assay. By contrast, the positional scanning strategy is used to define the most preferred residue or residues at each position of the sequence. All possible combinations of amino acids with activity exceeding some threshold level are then prepared and tested to identify explicit sequences with potent activities.[143,144]

Using competition ELISA experiments to map the linear epitopes recognized by two different monoclonal antibodies, both screening strategies converged to equivalent peptide sequences.[134,142,143] Similarly, either dual fixed residue or positional scanning screening of these hexapeptide libraries rediscovered the methionine and leucine enkephalin peptide families as competitive inhibitors of the binding of the μ-opioid-specific ligand [^{3}H]DAMGO to crude rat brain homogenates.[144,145] As the presence of a positively charged N-terminal amino group is generally regarded as a critical determinant for high-affinity peptide binding at the opioid receptors, the recent finding that potent ligands can identified from an N-terminally acetylated hexapeptide library is of particular interest.[146] The peptides Ac-RFMWMT-NH_2, Ac-RFMWMK-NH_2, and Ac-RFMWMR-NH_2 have low nanomolar affinities for the μ-, δ-, and κ_3- receptors and behave as potent μ-antagonists in the guinea pig ileum assay. Removal of the N-terminal acetyl group from these peptides results in a 500–1000-fold decrease in affinity for μ-receptor, and the compounds have been termed "acetalins" to reflect the critical contribution of the N-terminus to the opioid binding characteristics.[146] Another noteworthy application of these soluble hexamer libraries includes the discovery of novel antimicrobial peptides with activities against both Gram-positive and Gram-negative bacteria by iterative microdilution assays.[134,142]

Inhibitors of the human immunodeficiency virus (HIV) protease have been identified by screening sets of equimolar peptide mixtures, together containing more than

240 000 soluble tetrapeptides.[147] The library incorporated the aspartyl protease transition-state mimetic statine at one position of the tetramer sequence, along with 22 other amino acid monomers. The preferred residue at each position in the sequence was identified through the standard iterative cycle of assay and library resynthesis and yielded the peptide Ac-Phe-Ile-Sta-D-Leu-NH_2 (IC_{50} value of 1.4 μM) as the most potent protease inhibitor. This compound provided the basis for further substitutional analysis and was elaborated to a pentapeptide analog with an IC_{50} value of 5 nM. Antibody epitopes have been mapped by another iterative assay strategy using competition ELISA in a way that attempts to minimize the total number of sublibraries that need be synthesized and screened.[129]

Zuckermann and co-workers have employed an automated robotic synthesizer to prepare equimolar peptide mixtures according to the split synthesis method.[148] A library of 361 different compounds formed by randomly incorporating 19 amino acids at each of two positions within a decapeptide known to bind an anti-gp120 monoclonal antibody was screened with this antibody in a two-step procedure. After establishing an amino acid preference at one of the degenerate positions by competition ELISA measurements, the optimal substitution at the other residue was identified by affinity purifying a 19-component peptide mixture on the antibody using gel-filtration chromatography. The affinity-selected peptides were identified by mass spectrometry and amino acid analysis, and the relative amounts of each component recovered in the purification correlated with the peptides' receptor-binding affinities. This strategy has also been applied to the analysis of a 512-component hexamer library incorporating a limited set of nonnatural amino acid substitutions.[149]

Peptide Libraries Tethered to Solid Supports. When the split synthesis method of library construction is performed with beaded resin equipped with a non-cleavable linker moiety, the resulting libraries consist of collections of particles, each bearing (ideally) a single peptide sequence. These libraries can be screened for interaction with soluble receptors bearing some reporter group, and individual beads that bind the target protein subjected to Edman microsequence analysis to determine the sequence of the corresponding peptide ligands.[133,150,151] Lam *et al.* have used this method to prepare a library of $\sim 2.5 \times 10^6$ pentapeptides from 19 L-amino acids. In one assay the library was screened for binding to an anti-β-endorphin mAb 3E7 that was conjugated to alkaline phosphatase, and beads that were stained by ELISA were selected with a micromanipulator and sequenced.[133,150] The majority of the deduced sequences were clearly enkephalin-related, and as soluble peptides, several had affinities for the mAb within the range of 15 nM to 10 μM. That lower affinity ($K_i > 10$ μM) enkephalin analogs could also be identified in this screen was attributed to the presence of a high density of ligands on the bead surface simulating a high concentration of the ligand in solution.[150] Although this explanation is probably partially accurate, it is likely that the bivalency of the antibody used in this assay permits the formation of high avidity, multivalent interactions between the receptor and tethered ligands with low intrinsic affinity for the mAb. This method of library synthesis and ligand identification has been subsequently adapted for screening soluble peptide mixtures. Incorporating multicleavable linker groups into the beads permits the release of equimolar quantities of ligand at independent steps, such that after the screening sufficient peptide is retained on each bead for Edman sequence analysis.[152]

The nature of the bead material used in the construction and screening of support-tethered libraries has an important impact on the utility of this display format in ligand discovery. The beads must be resistant to all the organic solvents and reagents used in ligand synthesis and yet be well-behaved in the aqueous media where protein receptor binding studies are conducted. In particular, the polymer support should have low nonspecific protein binding characteristics to maximize the opportunity of detecting specific receptor interactions with the immobilized ligands. Some workers have preferred to use receptors that are directly radiolabeled in their binding studies with bead-libraries to minimize the background signals associated with enzyme-linked immunoassays.[151] The choice of the linker or spacer moiety on the bead from which the compounds are synthesized can also influence the accessibility of the ligands to macromolecular receptors in solution. Poly(dimethylacrylamide) resins derivatized with ethylenediamine, aminocaproic acid and β-alanine spacers have been successfully used in several studies.[133,150] Recently, a polyethylene-grafted cross-linked polystyrene resin termed TentaGel has been made commercially available by Rapp Polymere (Tubingen, Germany) for use as a carrier in tethered peptide library construction. The hydrophilic poly(oxyethylene) spacer moiety is equally well solvated in organic and aqueous solvents, and the relatively monodisperse bead material is available in a variety of diameters and with several different reactive functional groups (e.g., NH_2, OH, SH, Br, CO_2H). The 90-μm-diameter NH_2 beads for peptide synthesis give rise to *ca.* 2.9×10^6 beads/g of resin with a loading of 80–100 pmol per bead. Since the detection limit in conventional phenylthiohydantoin (PTH)-based Edman peptide microsequence analysis is *ca.* 1–5 pmol, each bead provides adequate material for direct ligand analysis after selection from a binding assay.

Encoded Combinatorial Synthesis. An important goal in screening diverse libraries of molecules produced through combinatorial synthesis is, of course, the identification of novel structures that interact with receptor targets of biological interest. For any given target protein, the probability of successfully identifying potent ligands through a process of randomly screening molecular repertoires in affinity-sensitive assays will undoubtedly increase as the size and structural diversity of the library is also increased. The generation and screening of very large combinatorial libraries assembled from an expanded set of molecular building blocks provides substantial challenges to the diversity technologies previously outlined in this section.

For example, to screen an immobilized synthetic peptide library containing 10^9 members constructed from the bead support described above would require treating kilograms of resin with liters of soluble receptor, clearly an impractical undertaking. A decrease in the bead diameter might permit a library of this diversity to occupy a manageable volume, but the quantity of peptide associated with any selected particle would quickly fall below the threshold required for microsequence analysis. The ability to dramatically expand the amino acid monomer set to

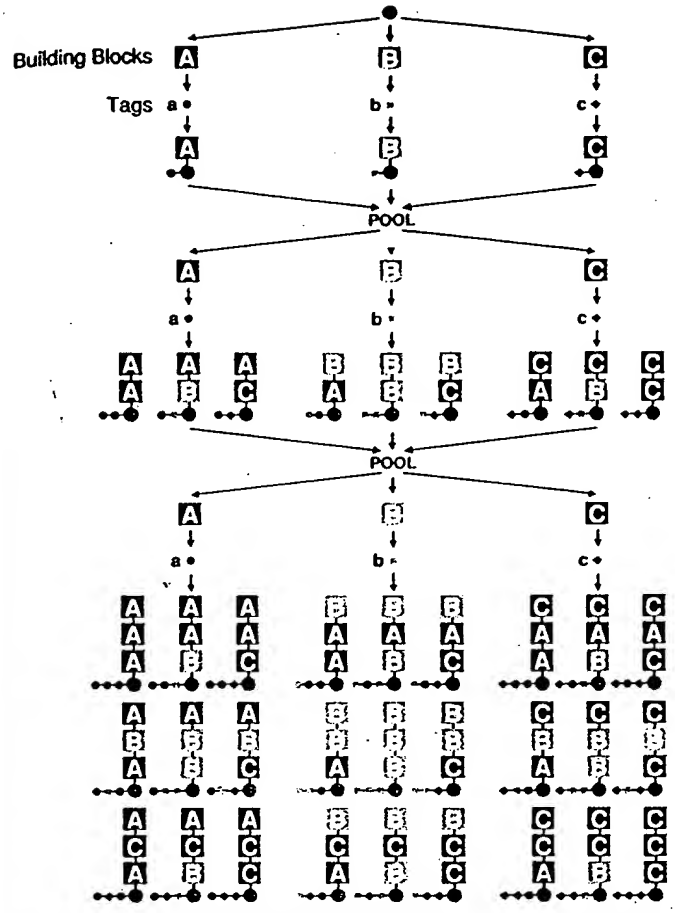


Figure 6. Schematic assembly of an encoded synthetic library.

include D-amino acids and nonnatural amino acids is also limited by the necessity of resolving chromatographically the PTH-derivatives of every building block used in the library. Moreover, direct ligand microsequencing remains an analysis option for peptides (Edman degradation) and oligonucleotides (Sanger dideoxy sequencing) only.

The iterative strategy of assaying compound mixtures and then synthesizing and retesting sublibraries avoids the problems associated with direct ligand structural analysis. However, this process can be very laborious, particularly with large libraries, since unique structural deconvolutions that lead to potent ligands are likely to be very rare. Furthermore, the complexity of nontethered libraries is limited in practice by the solubility of the mixtures, since the concentrations of individual components in the pools eventually become too small to exert any independent measurable effect on receptor binding or function.

In attempting to address some of these limitations, it was recognized that the products of a combinatorial synthesis on resin beads could be explicitly specified if it were possible to attach an identifier tag to the beads coincident with each monomer coupling step in the synthesis.[153] Each tag would then convey which monomer was coupled in a particular step of the synthesis, and the overall structure of a ligand on any bead could be deduced by reading the set of tags on that bead. An encoded version of the combinatorial synthesis previously outlined in Figure 2 is illustrated schematically in Figure 6.

Ideally, such tags should have a high information content, be amenable to very high sensitivity detection and decoding, and must be stable to reagents used in the ligand synthesis. A method using single-stranded oligonucleotides to encode combinatorial peptide syntheses on 10-μm-diameter polystyrene beads has recently been developed.[154] Peptides and nucleotides are assembled in parallel, alternating syntheses so that each bead bears many copies of both a single peptide sequence and a unique oligonucleotide identifier tag. The average stoichiometry of peptide to oligonucleotide per particle is readily controlled by varying the ratio of differentially functionalized linker residues coupled to the beads and can be heavily weighted in favor of the peptide molecules. Each amino acid monomer used in the synthesis is assigned a distinct contiguous nucleotide sequence or "codon" and hence the structure of the peptide assembled on any bead is reflected in the oligonucleotide sequence of the corresponding tag. The oligonucleotides share common 5′- and 3′-PCR priming sites and thus the beads can serve as templates for the PCR.

An oligonucleotide-encoded synthetic library of approximately 8.2×10^5 heptapeptides was screened for binding to an anti-peptide monoclonal antibody using a fluorescence activated cell sorting (FACS) instrument to select individual beads that strongly bind the antibody.[154] After PCR amplification of the oligonucleotide tags on sorted beads, the DNA was sequenced to determine the identity of the peptide ligands. This experiment clearly defined a consensus peptide sequence recognized by the mAb, and the affinities of peptides identified in this assay for the antibody ranged from 0.3 nM to 1.4 μM.

The concept of an oligonucleotide-encoded chemical synthesis has also been proposed independently by Brenner and Lerner.[155,156] They envisaged screening a library of soluble ligand–oligonucleotide hybrid molecules for receptor binding and proposed that the nucleotide tag could serve both as a coding device and as a handle for selectively enriching active ligands from the pool. More recently this group has disclosed the design of an orthogonally-protected polyfunctional linker group that can be used for the parallel, alternating solid-phase synthesis of peptides and oligonucleotides. They report that soluble peptide–oligonucleotide chimeras generated on this matrix can bind as anticipated to an anti-peptide antibody, as well as functioning as templates for amplification by the PCR.[157]

Zuckermann and co-workers have shown that by starting with an orthogonally differentiated diamine linker, parallel combinatorial synthesis can be used to generate a library of soluble chimeric peptides comprising a "binding" strand and a "coding" strand.[158] The coupling of either natural or nonnatural amino acid monomers to the binding strand was recorded by building an amino acid code comprised of four L-amino acids on the "coding" strand. Compounds were selected from equimolar peptide mixtures by affinity purification on a receptor and were resolved by HPLC. The sequence of the coding strand of individual purified molecules was then determined by Edman degradation to reveal the structure of the binding strand. An analogous peptidic coding scheme has also been recently described in detail by Nikolaiev *et al.*[159] With the appropriate choice of linker residues, either the binding or coding sequences or both can be cleaved from the resin. Thus it is possible to screen for receptor binding to tethered and to soluble ligands using this technique and to separate the contributions of the ligand and coding sequences to receptor interaction.

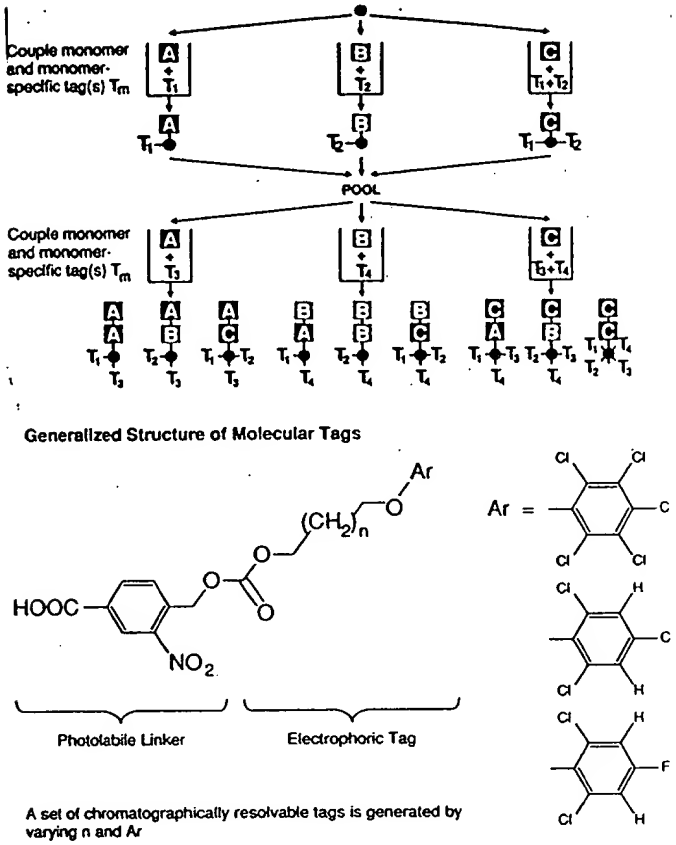


Figure 7. Binary encoded synthesis using gas chromatographically detectable chemically inert tags.

Still and colleagues[160] have disclosed an alternate approach to encoding combinatorial syntheses on resin beads wherein electron capture capillary gas chromatography (ECGC) is used in the analysis of the coding moieties. This scheme employs a series of chromatographically resolvable halocarbon derivatives as molecular tags which, when appended to reactive groups on the bead surface, can constitute a binary code that reflects the chemical history of any member of a library. In contrast to the oligonucleotide or peptide-coding approaches where the order of assembly of the chemical building blocks for any library member is preserved in the sequence of a single cognate tagging molecule, the binary strategy uses a uniquely defined mixture of tags to represent each building block at each particular step of the synthesis. Thus a set of N distinguishable tags can be used to encode the combinatorial synthesis of a library containing as many as 2^N different members.

The Still group has synthesized peptide libraries on 50–80-μm-diameter beads and determined the structures of ligands for a monoclonal antibody[160] and synthetic receptor[161] through the use of this coding strategy. The library construction follows the "split synthesis" protocol except that prior to each amino acid coupling, a substoichiometric addition of a tag mixture is made to the beads so that ~0.5% of the growing ligand chain is terminated for each tag added. The tags are coupled to the beads via a photolabile carbonate linker that can be cleaved upon irradiation at 365 nm (see Figure 7). After assembly, the library is screened for interaction with labeled receptor and stained beads are selected for analysis. Photolysis of individual beads in ~1 μL of DMF releases the haloarene tags for rapid analysis by ECGC. This technology seems well suited for analyzing combinatorial libraries comprised of up to ~10^6 members, although the sensitivity of electron capture detection may ultimately enable accurate sequence information to be obtained from beads of even smaller size than those used in this work. The chemically inert nature of the halocarbon tags is a particularly attractive feature of this coding approach, and should facilitate the screening of nonsequenceable organic molecules prepared by multistep combinatorial synthesis.

By contrast, encoding a combinatorial synthesis with oligonucleotides offers the advantage that tremendous levels of tag amplification can be achieved through the polymerase chain reaction. Thus it is possible to work with tiny quantities of DNA template and hence to use solid supports of microscopic dimensions in the syntheses. For example, the 10-μm-diameter beads described above give rise to 5×10^9 beads/g with a maximum ligand loading of ~20 fmol/bead. By using a combination of magnetic-activated and fluorescence-activated bead sorting techniques together with DNA sequence analysis, it should be practical to screen libraries containing as many as 10^8–10^9 synthetic compounds in soluble receptor binding assays. It seems likely that constraints on the sensitivity and/or throughput of other analytical procedures will ultimately restrict the scope of the nonamplifiable coding approaches to analyzing libraries of more limited diversity.

Acknowledgment. We thank Drs. A. Zaffaroni and G. Ringold for their interest, insights, and support in undertaking the application of combinatorial technologies to drug discovery and Mrs. Terri Nebozuk for help in preparing this manuscript.

References

(1) Hirschmann, R. Medicinal Chemistry in the Golden Age of Biology: Lessons from Steroid and Peptide Research. *Angew. Chem.* 1991, *30*, 1278–1301.
(2) Patchett, A. A. Excursions in Drug Discovery. *J. Med. Chem.* 1993, *36*, 2051–2058.
(3) Chabala, J. C. Pharmaceutical Manufacturers Association, Drug Discovery Management Subsection, September 19–21, 1993, Philadelphia, PA.
(4) Dower, W. J.; Fodor, S. P. A. The Search for Molecular Diversity (II) Recombinant and Synthetic Randomized Peptide Libraries *Annu. Rep. Med. Chem.* 1991, *26*, 271–280.
(5) Moos, W. H.; Green, G. D.; Pavia, M. R. Recent Advances in the Generation of Molecular Diversity. *Annu. Rep. Med. Chem.* 1993 *28*, 315–324.
(6) Zuckermann, R. N. The Chemical Synthesis of Peptidomimeti Libraries. *Curr. Opin. Struct. Biol.* 1993, *3*, 580–584.
(7) Pavia, M. R., Sawyer, T. K., Moos, W. H., Eds. The Generation o Molecular Diversity. *Bioorg. Med. Chem. Lett.* 1993, *3*, 381–475
(8) Zoller, M. J.; Smith, M. Oligonucleotide-directed mutagenesis usin M13-derived vectors: an efficient and general procedure for th production of point mutations in any fragment of DNA. *Nuclei Acids Res.* 1982, *10*, 6487–6500.
(9) Dalbadie-McFarland, G.; Cohen, L. W.; Riggs, A. D.; Morin, C Itakura, K.; Richards, J. H. Oligonucleotide-directed mutagenesi as a general and powerful method for studies of protein functior *Proc. Natl. Acad. Sci. U.S.A.* 1982, *79*, 6409–6413.
(10) Matteucci, M. D.; Heyneker, H. L. Targeted random mutagenesi the use of ambiguously synthesized oligonucleotides to mutageni sequences immediately 5′ of an ATG initiation codon. *Nuclei Acids Res.* 1983, *11*, 3113–3121.
(11) Wells, J. A.; Vasser, M.; Powers, D. B. Cassette mutagenesis: e efficient method for generation of multiple mutations at define sites. *Gene* 1985, *34*, 315–323.
(12) Horwitz, M. S. Z.; Loeb, L. A. Promoters Selected from Rando DNA Sequences. *Proc. Natl. Acad. Sci. U.S.A.* 1986, *83*, 7405 7409.
(13) Reidhaar-Olson, J. F.; Sauer, R. T. Combinatorial Cassette M tagenesis as a Probe of the Informational Content of Prote Sequences. *Science* 1988, *241*, 53–57.
(14) Young, R. A.; Davis, R. W. Yeast RNA polymerase II genes: isolati with antibody probes. *Science* 1983, *222*, 778–782.
(15) Young, R. A.; Davis, R. W. Efficient isolation of genes by usi antibody probes. *Proc. Natl. Acad. Sci. U.S.A.* 1983, *80*, 119 1198.

(16) Aruffo, A.; Seed, B. Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. *Proc. Natl. Acad. Sci. U.S.A.* 1987, *84*, 8573–8577.
(17) Seed, B.; Aruffo, A. Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc. Natl. Acad. Sci. U.S.A.* 1987, *84*, 3365–3369.
(18) Wagstrom, C.; Cwirla, S.; Dower, W. J. Unpublished results.
(19) Charbit, A.; Boulain, J. C.; Ryter, A.; Hofnung, M. Probing the topology of a bacterial membrane protein by genetic insertion of a foreign epitope; expression at the cell surface. *EMBO J.* 1986, *5*, 3029–3037.
(20) Schorr, J.; Kanpp, B.; Hundt, E.; Kupper, H.; Amann, E. Surface expression of malarial antigens in *E. coli* and *S. typhimurium*: induction of seum antibody response upon oral vaccination of mice. *Vaccines 91*; Chanock, R. M., Ginsberg, H. S., Brown, F., Lerner, R. A., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1991; pp 387–392.
(21) Agterberg, M.; Adriaanse, H.; van Bruggen, A.; Karperien, M.; Tommassen, J. Outer-membrane PhoE protein of *E. coli* K-12 as an exposure vector: possibilities and limitations. *Gene* 1990, *88*, 37–45.
(22) Fuchs, P.; Breitling, F.; Dubel, S.; Seehaus, T.; Little, M. Targeting recombinant antibodies to the surface of *Escherichia coli*: fusion to a peptidoglycan associated lipoprotein. *Bio/Tech.* 1991, *9*, 1369–1372.
(23) Thiry, G.; Clippe, A.; Scarcez, T.; Petre, J. Cloning of DNA sequences encoding foreign peptides and their expression in the K88 pili. *Appl. Environ. Microbiol.* 1989, *55*, 984–993.
(24) Kuwajima, G.; Asaka, J.-I.; Fujiwara, T.; Fujiwara, T.; Nakano, K.; Kondoh, E. Presentation of an antigenic determinant from hen egg-white lysozyme on the flagellar filament of *Escherichia coli*. *Bio/Tech.* 1988, *6*, 1080–1083.
(25) Hansson, M.; Stahl, S.; Nguyen, T. N.; Bachi, T.; Robert, A.; Binz, H.; Sjolander, A.; Uhlen, M. Expression of recombinant proteins on the surface of the coagulase-negative bacterium *Staphylococcus xylosus*. *J. Bacteriol.* 1992, *174*, 4239–4245.
(26) Klauser, T.; Pohler, J.; Meyer, T. F. Extracellular transport of cholera toxin B subunit using *Neisseria* IgA protease β-domain: conformation-dependent outer membrane translocation. *EMBO J.* 1990, *9*, 1991–1999.
(27) Brown, S. Engineered iron oxide-adhesion mutants of the *Escherichia coli* phage λ receptor. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 8651–8655.
(28) Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 1985, *228*, 1315–1317.
(29) Parmley, S. F.; Smith, G. P. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* 1988, *73*, 305–318.
(30) Cwirla, S.; Peters, E. A.; Barrett, R. W.; Dower, W. J. Peptides on phage: a vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. U.S.A.* 1990, *87*, 6378–6382.
(31) Scott, J. K.; Smith, G. P. Searching for peptide ligands with an epitope library. *Science* 1990, *249*, 386–390.
(32) Devlin, J. J.; Panganiban, L. C.; Devlin, P. E. Random peptide libraries: a source of specific protein binding molecules. *Science* 1990, *249*, 404–406.
(33) Il'ichev, A. A.; Minenkova, O. O.; Tat'kov, S. I.; Karpyshev, N. N.; Eroshkin, A. M.; Petrenko, V. A.; Sandakhchiev, L. S. Production of a viable variant of the M13 phage with a foreign peptide inserted into the basic coat protein. *Doklady Akademii Nauk. S.S.S.R.* 1989, *307*, 481–483.
(34) Greenwood, J.; Willis, A. E.; Perham, R. N. Multiple display of foreign peptides on a filamentous bacteriophage. *J. Mol. Biol.* 1991, *220*, 821–827.
(35) Peters, E. A.; Cwirla, S.; Dower, W. J. Unpublished results.
(36) Felici, F.; Castagnoli, L.; Musacchio, A.; Jappelli, R.; Cesareni, G. Selection of Antibody Ligands from a Large Library of Oligopeptides Expressed on a Multivalent Exposition Vector. *J. Mol. Biol.* 1991, *222*, 301–310.
(37) Kang, A. S.; Barbas, C. F.; Janda, K. K.; Benkovic, S. J.; Lerner, R. A. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc. Natl. Acad. Sci. U.S.A.* 1991, *88*, 4363–4366.
(38) McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* 1990, *348*, 552–554.
(39) Hoogenboom, H. R.; Griffiths, A. D.; Johnson, K. S.; Chiswell, D. J.; Hudson, P.; Winter, G. Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Res.* 1991, *19*, 4133–4137.
(40) Marks, J. D.; Hoogenboom, H. R.; Bonnert, T. P.; McCafferty, J.; Griffiths, A. D.; Winter, G. By-Passing Immunization: Human Antibodies from V-Gene Libraries Displayed on Phage. *J. Mol. Biol.* 1991, *222*, 581.
(41) Barbas, C. F., III; Bain, J. D.; Hoekstra, D. M.; Lerner, R. A. Semisynthetic Combinatorial Antibody Libraries: A Chemical Solution to the Diversity Problem. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 4457–4461.
(42) Hoogenboom, H. R.; Marks, J. D.; Griffiths, A. D.; Winter, G. Building antibodies from their genes. *Immunol. Rev.* 1992, *130*, 41–68.
(43) Soderlind, E.; Simonsson, A. C.; Borrebaeck, C. A. K. Phage display technology in antibody engineering: Design of phagemid vectors and *in vitro* maturation systems. *Immunol. Rev.* 1992, *130*, 109–124.
(44) Bass, S.; Green, R.; Wells, J. A. Hormone Phage: An enrichment method for variant proteins with altered binding properties. *Proteins: Struct. Funct. Genet.* 1990, *8*, 309–314.
(45) Lowman, H. B.; Bass, S. H.; Simpson, N.; Wells, J. A. Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry* 1991, *30*, 10832–10838.
(46) Markland, W.; Roberts, B. L.; Saxena, M. J.; Guterman, S. K.; Ladner, R. C. Design, construction and function of a multicopy display vector using fusions to the major coat protein of bacteriophage M13. *Gene* 1991, *109*, 13–19.
(47) Roberts, B. L.; Markland, W.; Ley, A. C.; Kent, R. B.; White, D. W.; Guterman, S. K.; Ladner, R. C. Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 2429–2433.
(48) Pannekoek, H.; van Meijer, M.; Schleef, R. R.; Loskutoff, D. J.; Barbas, C. F., III. Functional display of human plasminogen-activator inhibitor 1 (PAI-1) on phages: novel perspectives for structure-function analysis by error-prone DNA synthesis. *Gene* 1993, *128*, 135–140.
(49) McCafferty, J.; Jackson, R. H.; Chiswell, D. J. Phage-enzymes: expression and affinity chromatography of functional alkaline phosphatase on the surface of bacteriophage. *Prot. Eng.* 1991, *4*, 955–961.
(50) Corey, D. R.; Shiau, A. K.; Yang, Q.; Janowski, B. A.; Craik, C. S. Trypsin display on the surface of bacteriophage. *Gene* 1993, *128*, 129–134.
(51) Scott, J. K. Identifying lead peptides from epitope libraries. *Drug Design*; Weiner, D. B., Williams, W. V., Eds.; CRC Press: Boca Raton, FL, 1993; pp 1–61.
(52) Dower, W. J. Phage power. *Curr. Biol.* 1992, *2*, 251–253.
(53) Stephen, C. W.; Lane, D. P. Mutant Conformation of p53: Precise epitope mapping using a filamentous phage epitope library. *J. Mol. Biol.* 1992, *225*, 577–583.
(54) Yayon, A.; Aviezer, D.; Safran, M.; Gross, J. L.; Heldman, Y.; Cabilly, S.; Givol, D.; Katchalski-Katzir, E. Isolation of peptides that inhibit binding of basic fibroblast growth factor to its receptor from a random phage-epitope library. *Proc. Natl. Acad. Sci. U.S.A.* 1993, *90*, 10643–10647.
(55) Felici, F.; Luzzago, A.; Folgori, A.; Cortese, R. Mimicking of discontinuous epitopes by phage-displayed peptides, II. Selection of clones recognized by a protective monoclonal antibody against the *Bordetella pertussis* toxin from phage peptide libraries. *Gene* 1993, *128*, 21–27.
(56) Luzzago, A.; Felici, F.; Tramontano, A.; Pessi, A.; Cortese, R. Mimicking of discontinuous epitopes by phage-displayed peptides, I. Epitope mapping of human H ferritin using a phage library of constrained peptides. *Gene* 1993, *128*, 51–57.
(57) Balass, M.; Heldman, Y.; Cabilly, S.; Givol, D.; Katchalski-Katzir, E.; Fuchs, S. Identification of a hexapeptide that mimics a conformation-dependent binding site of acetylcholine receptor by use of a phage-epitope library. *Proc. Natl. Acad. Sci. U.S.A.* 1993, *90*, 10638–10642.
(58) Hoess, R.; Brinkmann, U.; Handel, T.; Pastan, I. Identification of a peptide which binds to the carbohydrate-specific monoclonal antibody B3. *Gene* 1993, *128*, 43–49.
(59) Kay, B. K.; Adey, N. B.; He, Y.-S.; Manfredi, J. P.; Mataragnon, A. H.; Fowlkes, D. M. An M13 phage library displaying random 38-amino-acid peptides as a source of novel sequences with affinity to selected targets. *Gene* 1993, *128*, 59–65.
(60) Oldenburg, K. R.; Loganathan, D.; Goldstein, I. J.; Schultz, P. G.; Gallop, M. A. Peptide ligands for a sugar-binding protein isolated from a random peptide library. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 5393–5397.
(61) Scott, J. K.; Loganathan, D.; Easley, R. B.; Gong, X.; Goldstein, I. J. A family of concanavalin A-binding peptides from a hexapeptide epitope library. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 5398–5402.
(62) Smith, G. P.; Schultz, D. A.; Ladbury, J. E. A Ribonuclease S-Peptide Antagonist Discovered with a Bacteriophage Display Library. *Gene* 1993, *128*, 37–42.
(63) O'Neil, K. T.; Hoess, R. H.; Jackson, S. A.; Ramachandran, N. S.; Mousa, S. A.; DeGrado, W. F. Identification of novel peptide antagonists for GPIIb/IIIa from a conformationally constrained phage peptide library. *Proteins: Struct. Funct. Genet.* 1992, *14*, 509–515.
(64) Koivunen, E.; Gay, D. A.; Ruoslahti, E. Selection of peptides binding to the $\alpha_5\beta_1$ integrin from a phage display library. *J. Biol. Chem.* 1993, *268*, 20205–20210.
(65) Renschler, M. F.; Bhatt, R. R.; Dower, W. J.; Levy, R. Synthetic peptide ligands of the immunoglobulin receptor induce programmed cell death in a human B-cell lymphoma. *Proc. Natl. Acad. Sci. U.S.A.* In press.

(66) Hammer, J.; Valsasnini, P.; Tolba, K.; Bolin, D.; Higelin, J.; Takacs, B.; Sinigaglia, F. Promiscuous and Allele-specific anchors in HLA-DR-binding peptides. *Cell* 1993, *74*, 197–203.
(67) Blond-Elguindi, S.; Cwirla, S. E.; Dower, W. J.; Lipshutz, R. J.; Sprang, S. R.; Sambrook, J. F.; Gething, M.-J. Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BIP. *Cell* 1993, *75*, 717–728.
(68) Matthews, D. J.; Wells, J. A. Substrate phage: selection of protease substrates by monovalent phage display. *Science* 1993, *260*, 1113–1117.
(69) Smith, M. M.; Shi, L.; Navre, M. Manuscript in preparation.
(70) Cull, M. G.; Miller, J. F.; Schatz, P. J. Screening for receptor ligands using large libraries of peptides linked to the C terminus of the *lac* repressor. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 1865–1869.
(71) Peters, E. A.; Schatz, P. J.; Dower, W. J. Submitted for publication.
(72) Schatz, P. J. Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: A 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. *Bio/Tech.* 1993, *11*, 1138–1143.
(73) Schatz, P. J.; Barrett, R. W. Unpublished results.
(74) Wagstrom, C.; Cwirla, S.; Dower, W. J. Unpublished results.
(75) Mattheakis, L. C.; Bhatt, R.; Dower, W. J. An *in vitro* polysome display system for identifying ligands from very large peptide libraries. Submitted.
(76) Barrett, R. W.; Cwirla, S. E.; Ackerman, M. S.; Olson, A. M.; Peters, E. A.; Dower, W. J. Selective enrichment and characterization of high affinity ligands from collections of random peptides on filamentous phage. *Anal. Biochem.* 1992, *204*, 357–364.
(77) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. A General Method for Site-Specific Incorporation of Unnatural Amino Acids into Proteins. *Science* 1989, *244*, 182–188.
(78) McNeil, J. B.; Smith, M. Saccharomyces cerevisiae CYC1 mRNA 5′-end positioning: analysis by in vitro mutagenesis, using synthetic duplexes random mismatch base pairs. *Mol. Cell Biol.* 1985, *5*, 3545–3551.
(79) Ner, S. S.; Goodin, D. B.; Smith, M. A simple and efficient procedure for generating random point mutations and for codon replacements using mixed oligodeoxynucleotides. *DNA* 1988, *7*, 127–134.
(80) Arkin, A.; Youvan, D. C. Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis. *Bio/Tech.* 1992, *10*, 297–300.
(81) *Comprehensive Medicinal Chemistry*; Hansch, C., Sammes, P. G., Taylor, J. B., Eds.; Pergamon Press: Oxford, 1990; Vol. 2.
(82) Jung, G.; Beck-Sickinger, A. G. Multiple Peptide Synthesis Methods and Their Applications. *Angew. Chem., Int. Ed. Engl.* 1992, *31*, 367–383.
(83) Merrifield, R. B. Solid Phase Peptide Synthesis I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* 1963, *85*, 2149–2154.
(84) Geysen, H. M.; Meloen, R. H.; Barteling, S. J. Use of Peptide Synthesis to Probe Viral Antigens for Epitopes to a Resolution of a Single Amino Acid. *Proc. Natl. Acad. Sci. U.S.A.* 1984, *81*, 3998–4002.
(85) Valerio, R. M.; Bray, A. M.; Campbell, R. A.; Dipasquale, A.; Margellis, C.; Rodda, S. J.; Geysen, H. M.; Maeji, N. J. Multipin Peptide Synthesis at the Micromole Scale Using 2-Hydroxyethyl Methacrylate Grafted Polyethylene Supports. *Int. J. Pept. Protein Res.* 1993, *42*, 1–9.
(86) Bray, A. M.; Maeji, N. J.; Geysen, H. M. The Simultaneous Multiple Production of Solution Phase Peptides: Assessment of the Geysen Method of Simultaneous Peptide Synthesis. *Tetrahedron Lett.* 1990, *31*, 5811–5814.
(87) Valerio, R. M.; Benstead, M.; Bray, A. M.; Campbell, R. A.; Maeji, N. J. Synthesis of Peptide Analogs Using the Multipin Peptide Synthesis Method. *Anal. Biochem.* 1991, *197*, 168–177.
(88) Bray, A. M.; Maeji, N. J.; Jhingran, A. G.; Valerio, R. M. Gas Phase Cleavage of Peptides from a Solid Support with Ammonia Vapor. Application in Simultaneous Multiple Peptide Synthesis. *Tetrahedron Lett.* 1991, *32*, 6163–6166.
(89) Geysen, H. M.; Rodda, S. J.; Mason, T. J.; Tribbick, G.; Schoofs, P. G. Strategies for Epitope Analysis Using Peptide Synthesis. *J. Immunol. Methods* 1987, *102*, 259–274.
(90) Geysen, H. M.; Rodda, S. J.; Mason, T. J. A Priori Delineation of a Peptide Which Mimics a Discontinuous Antigenic Determinant. *Mol. Immunol.* 1986, *23*, 709–715.
(91) Geysen, H. M.; Barteling, S. J.; Meloen, R. H. Small Peptides Induce Antibodies with a Sequence and Structural Requirement for Binding Antigen Comparable to Antibodies Raised Against the Native Protein. *Proc. Natl. Acad. Sci. U.S.A.* 1985, *82*, 178–182.
(92) Middeldorp, J. M.; Meloen, R. H. Epitope-Mapping on the Epstein-Barr Virus Major Capsid Protein Using Systematic Synthesis of Overlapping Peptides. *J. Virol. Methods* 1988, *21*, 147–159.
(93) Gombert, F.; Blecha, W.; Tahtinen, M.; Ranki, A.; Pfeiffer, S.; Troger, W.; Braun, R.; Muller-Lantzsch, N.; Jung, G.; Rubsamen-Waigmann, H.; Krohn, K. Antigenic Epitopes of NEF Proteins from Different HIV-1 Strains as Recognized by Sera from Patients with Manifest and Latent HIV Infection. *Virology* 1990, *176*, 458–466.
(94) Maeji, N. J.; Bray, A. M.; Geysen, H. M. Multi-Pan Peptide Synthesis Strategy for T Cell Determinant Analysis. *J. Immunol. Methods* 1990, *134*, 23–33.
(95) Gammon, G.; Geysen, H. M.; Apple, R. J.; Pickett, E.; Palmer, M.; Ametani, A.; Sercarz, E. E. T Cell Determinant Structure: Cores and Determinant Envelopes in Three Mouse Major Histocompatibility Complex Haplotypes. *J. Exp. Med.* 1991, *173*, 609–617.
(96) Mutch, D. A.; Rodda, S. J.; Benstead, M.; Valerio, R. M.; Geysen, H. M. Effects of End Groups on the Stimulatory Capacity of Minimal Length T Cell Determinant Peptides. *Pept. Res.* 1991, *4*, 132–137.
(97) Wang, J.-x.; Bray, A. M.; DiPasquale, A. J.; Maeji, N. J.; Geysen, H. M. Application of the Multipin Peptide Synthesis Technique for Peptide Receptor Binding Studies: Substance P as a Model System. *Bioorg. Med. Chem. Lett.* 1993, *3*, 447–450.
(98) Wang, J.-x.; DiPasquale, A. J.; Bray, A. M.; Maeji, N. J.; Geysen, H. M. Study of Stereo-Requirements of Substance P Binding to NK1 Receptors Using Analogs with Systematic D-Amino Acid Replacements. *Bioorg. Med. Chem. Lett.* 1993, *3*, 451–456.
(99) Spellmeyer, D. C.; Brown, S.; Stauber, G. B.; Geysen, H. M.; Valerio, R. Endothelin Receptor Ligands. Replacement Net Approach to SAR Determination of Potent Hexapeptides. *Bioorg. Med. Chem. Lett.* 1993, *3*, 519–524.
(100) Spellmeyer, D. C.; Brown, S.; Stauber, G. B.; Geysen, H. M.; Valerio, R. Endothelin Receptor Ligands. Multiple D-Amino Acid Replacement Net Approach. *Bioorg. Med. Chem. Lett.* 1993, *3*, 1253–1256.
(101) Houghten, R. A. General Method for the Rapid Solid-Phase Synthesis of Large Numbers of Peptides: Specificity of Antigen-Antibody Interaction at the Level of Individual Amino Acids. *Proc. Natl. Acad. Sci. U.S.A.* 1985, *82*, 5131–5135.
(102) Houghten, R. A.; Bray, M. K.; Degraw, S. T.; Kirby, C. J. Simplified Procedure For Carrying Out Simultaneous Multiple Hydrogen Fluoride Cleavages of Protected Peptide Resins. *Int. J. Pept. Protein Res.* 1986, *27*, 673–678.
(103) Beck-Sickinger, A. G.; Durr, H.; Jung, G. Semi-Automated T-Bag Peptide Synthesis Using 9-Fluorenylmethoxycarbonyl Strategy and Benzotriazol-1-yl-tetramethyluronium Tetrafluoroborate Activation. *Pept. Res.* 1991, *4*, 88–94.
(104) Beck-Sickinger, A. G.; Gaida, W.; Schnorrenberg, G.; Lang, R.; Jung, G. Neuropeptide Y: Identification of the Binding Site. *Int. J. Pept. Protein Res.* 1990, *36*, 522–530.
(105) Beck-Sickinger, A. G.; Jung, G.; Gaida, W.; Koppen, H.; Schnorrenberg, G.; Lang, R. Structure/Activity Relationships of C-Terminal Neuropeptide Y Peptide Segments and Analogs Composed of Sequence 1–4 Linked to 25–36. *Eur. J. Biochem.* 1990, *194*, 449–456.
(106) Zimmermann, N.; Beck-Sickinger, A. G.; Folkers, G.; Krickl, S.; Muller, I.; Jung, G. Conformational and Epitope Mapping of Herpes-Simplex-Virus Type-1 Thymidine Kinase Using Synthetic Peptide Segments. *Eur. J. Biochem.* 1991, *200*, 519–528.
(107) Ruggeri, Z. M.; Houghten, R. A.; Russell, S. R.; Zimmerman, T. S. Inhibition of Platelet Function with Synthetic Peptides Designed to be High-Affinity Antagonists of Fibrinogen Binding to Platelets. *Proc. Natl. Acad. Sci. U.S.A.* 1986, *83*, 5708–5712.
(108) Tjoeng, F. S.; Towery, D. S.; Bulock, J. W.; Whipple, D. E.; Fok, K. F.; Williams, M. H.; Zupec, M. E.; Adams, S. P. Multiple Peptide Synthesis Using a Single Solid Support (MPS3). *Int. J. Pept. Protein Res.* 1990, *35*, 141–146.
(109) Rutter, W. J.; Santi, D. V. General Method for Producing and Selecting Peptides With Specific Properties. U.S. Pat. 5,010,175, 1991.
(110) Birkett, A. J.; Soler, D. F.; Wolz, R. L.; Bond, J. S.; Wiseman, J.; Berman, J.; Harris, R. B. Determination of Enzyme Specificity in a Complex Mixture of Peptide Substrates by N-Terminal Sequence Analysis. *Anal. Biochem.* 1991, *196*, 137–143.
(111) Petithory, J. R.; Masiarz, F. R.; Kirsch, J. F.; Santi, D. V.; Malcolm, B. A. A Rapid Method for Determination of Endoproteinase Substrate Specificity: Specificity of the 3C Proteinase from Hepatitis A Virus. *Proc. Natl. Acad. Sci. U.S.A.* 1991, *88*, 11510–11514.
(112) Blankemeyer-Menge, B.; Frank, R. Simultaneous Multiple Synthesis of Protected Peptide Fragments on 'Allyl'-Functionalized Cellulose Disk Supports. *Tetrahedron Lett.* 1988, *29*, 5871–5874.
(113) Frank, R.; Doring, R. Simultaneous Multiple Peptide Synthesis Under Continuous Flow Conditions on Cellulose Paper Disks as Segmental Solid Supports. *Tetrahedron* 1988, *44*, 6031–6040.
(114) Eichler, J.; Beyermann, M.; Bienert, M. Application of Cellulose Paper as Support Material in Simultaneous Solid-Phase Peptide Synthesis. *Collect. Czech. Chem. Commun.* 1989, *54*, 1746–1752.
(115) Frank, R. Strategies and Techniques in Simultaneous Multiple Peptide Synthesis Based on the Segmentation of Membrane Type Supports. *Bioorg. Med. Chem. Lett.* 1993, *3*, 425–430.
(116) Eichler, J.; Bienert, M.; Stierandova, A.; Lebl, M. Evaluation of Cotton as a Carrier for Solid-Phase Peptide Synthesis. *Pept. Res.* 1991, *4*, 296–307.
(117) Schmidt, M.; Eichler, J.; Odarjuk, J.; Krause, E.; Beyermann, M.; Bienert, M. Multiple Peptide Synthesis Using Cellulose-Based Carriers: Synthesis of Substance P Diastereomers and their Histamine-Releasing Activity. *Bioorg. Med. Chem. Lett.* 1993, *3*, 441–446.

(118) Frank, R. Spot Synthesis: An Easy Technique for the Positionally Addressable, Parallel Chemical Synthesis on a Membrane Support. *Tetrahedron* 1992, *48*, 9217–9232.
(119) Berg, R. H.; Almdal, K.; Batsberg Pedersen, W.; Holm, A.; Tam, J. P.; Merrifield, R. B. Long-Chain Polystyrene-Grafted Polyethylene Film Matrix: A New Support for Solid-Phase Peptide Synthesis. *J. Am. Chem. Soc.* 1989, *111*, 8024–8026.
(120) Schnorrenberg, G.; Gerhardt, H. Fully Automatic Simultaneous Multiple Peptide Synthesis in Micromolar Scale-Rapid Synthesis of Series of Peptides for Screening in Biological Assays. *Tetrahedron* 1989, *45*, 7759–7764.
(121) Gausepohl, H.; Boulin, C.; Kraft, M.; Frank, R. W. Automated Multiple Peptide Synthesis. *Pept. Res.* 1992, *5*, 315–320.
(122) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C. Design, Construction and Application of a Fully Automated Equimolar Peptide Mixture Synthesizer. *Int. J. Pept. Protein Res.* 1992, *40*, 497–506.
(123) Dryland, A.; Sheppard, R. C. Peptide Synthesis. Part 8. A System for Solid-Phase Synthesis Under Low Pressure Continuous Flow Conditions. *J. Chem. Soc., Perkin Trans. 1* 1986, 125–137.
(124) Krchnak, V.; Vagner, J.; Mach, O. Multiple Continuous-Flow Solid-Phase Peptide Synthesis. Synthesis of an HIV Antigenic Peptide and its Omission Analogs. *Int. J. Pept. Protein Res.* 1989, *33*, 209–213.
(125) Meldal, M.; Bisgaard Holm, C.; Bojesen, G.; Havsteen Jakobsen, M.; Holm, A. Multiple Column Peptide Synthesis, Part 2. *Int. J. Pept. Protein Res.* 1993, *41*, 250–260.
(126) Zuckermann, R. N.; Banville, S. C. Automated Peptide-Resin Deprotection Cleavage by a Robotic Workstation. *Pept. Res.* 1992, *5*, 169–174.
(127) Neimark, J.; Briand, J.-P. Development of a Fully Automated Multichannel Peptide Synthesizer with Integrated TFA Cleavage Capability. *Pept. Res.* 1993, *6*, 219–228.
(128) Gras-Masse, H.; Ameisen, J.-C.; Boutillon, C.; Rouaix, F.; Bossus, M.; Deprez, B.; Capron, A.; Tartar, A. Synthetic Vaccines: The Mixotope Strategy. *Peptides: Chemistry, Structure and Biology, Proceedings of the 12th American Peptide Symposium*; Smith, J. A., Rivier, J. E., Eds.; Escom: Leiden, 1992; pp 842–844.
(129) Blake, J.; Litzi-Davis, L. Evaluation of Peptide Libraries: An Iterative Strategy To Analyze the Reactivity of Peptide Mixture with Antibodies. *Bioconj. Chem.* 1992, *3*, 510–513.
(130) Furka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G. 1988, Abstr. 14th Int. Congr. Biochem., Prague, Czechoslovakia, Vol. 5, p 47. Abstr. 10th Int. Symp. Med. Chem., Budapest, Hungary, p 288.
(131) Furka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G. General Method for Rapid Synthesis of Multicomponent Peptide Mixtures. *Int. J. Pept. Protein Res.* 1991, *37*, 487–493.
(132) Sebestyen, F.; Dibo, G.; Kovacs, A.; Furka, A. Chemical Synthesis of Peptide Libraries. *Bioorg. Med. Chem. Lett.* 1993, *3*, 413–418.
(133) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. A New Type of Synthetic Peptide Library for Identifying Ligand-Binding Activity. *Nature* 1991, *354*, 82–84.
(134) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. Generation and Use of Synthetic Peptide Combinatorial Libraries for Basic Research and Drug Discovery. *Nature* 1991, *354*, 84–86.
(135) Geysen, H. M.; Mason, T. J. Screening Chemically Synthesized Peptide Libraries for Biologically Relevant Molecules. *Bioorg. Med. Chem. Lett.* 1993, *3*, 397–404.
(136) Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. Light-Directed, Spatially Addressable Parallel Chemical Synthesis. *Science* 1991, *251*, 767.
(137) Pirrung, M. C.; Read, J. L.; Fodor, S. P. A.; Stryer, L. U.S. Pat. 5,143,854 1992.
(138) Jacobs, J. W.; Fodor, S. P. A. Combinatorial Chemistry: Applications of Light Directed Chemical Synthesis. *Trends Biotechnol.* 1994, *12*, 19–26.
(139) Holmes, C. P.; Adams, C. L.; Fodor, S. P. A.; Yu-Yang, P. New Techniques in Random Screening: The use of Light-Directed Spatially Addressable Chemical Synthesis. In *Perspectives in Medicinal Chemistry*; Testa, B., Kyburz, E., Fuher, W., Giger, R., Eds.; VHCA: Basel, 1992; pp 489–500.
(140) Adams, C. L.; Kochersperger, L. M.; Mortensen, R. B.; Aldwin, L. A.; Holmes, C. P. New Methodologies in the Preparation of Peptide Libraries. In *Techniques in Protein Chemistry V*; Crabb, J. W., Ed.; Academic Press: San Diego, 1994; pp 525–532.
(141) Holmes, C. P.; Fodor, S. P. A. Photolithography and Combinatorial Chemistry in Screening. In *Innovation and Perspectives in Solid Phase Synthesis and Complementary Technologies*; Epton, R., Ed.; I Intercept, UK, 1994, in press.
(142) Houghten, R. A.; Appel, J. R.; Blondelle, S. E.; Cuervo, J. H.; Dooley, C. T.; Pinilla, C. The Use of Synthetic Peptide Combinatorial Libraries for the Identification of Bioactive Peptides. *BioTechniques* 1992, *13*, 412–421.
(143) Pinilla, C.; Appel, J. R.; Blanc, P.; Houghten, R. A. Rapid Identification of High Affinity Peptide Ligands Using Positional Scanning Synthetic Peptide Combinatorial Libraries. *BioTechniques* 1992, *13*, 901–905.
(144) Dooley, C. T.; Houghten, R. A. The Use of Positional Scanning Synthetic Peptide Combinatorial Libraries for the Rapid Determination of Opioid Receptor Ligands. *Life Sci.* 1993, *52*, 1509–1517.
(145) Houghten, R. A.; Dooley, C. T. The Use of Synthetic Peptide Combinatorial Libraries for the Determination of Peptide Ligands in Radio-Receptor Assays: Opioid Peptides. *Bioorg. Med. Chem. Lett.* 1993, *3*, 405–412.
(146) Dooley, C. T.; Chung, N. N.; Schiller, P. W.; Houghten, R. A. Acetalins: Opioid Receptor Antagonists Determined Through the Use of Synthetic Peptide Combinatorial Libraries. *Proc. Natl. Acad. Sci. U.S.A.* 1993, *90*, 10811–10815.
(147) Owens, R. A.; Gesellchen, P. D.; Houchins, B. J.; DiMarchi, R. D. The Rapid Identification of HIV Protease Inhibitors Through the Synthesis and Screening of Defined Peptide Mixtures. *Biochem. Biophys. Res. Commun.* 1991, *181*, 402–408.
(148) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C.; Santi, D. V. Identification of Highest-Affinity Ligands by Affinity Selection from Equimolar Peptide Mixtures Generated by Robotic Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 4505–4509.
(149) Kerr, J. M.; Banville, S. C.; Zuckermann, R. N. Identification of Antibody Mimotopes Containing Non-Natural Amino Acids by Recombinant and Synthetic Peptide Library Affinity Selection Methods. *Bioorg. Med. Chem. Lett.* 1993, *3*, 463–468.
(150) Lam, K. S.; Hruby, V. J.; Lebl, M.; Knapp, R. J.; Kazmierski, W. M.; Hersh, E. M.; Salmon, S. E. The Chemical Synthesis of Large Random Peptide Libraries and Their Use for the Discovery of Ligands for Macromolecular Acceptors. *Bioorg. Med. Chem. Lett.* 1993, *3*, 419–424.
(151) Kassarjian, A.; Schellenberger, V.; Turck, C. W. Screening of Synthetic Peptide Libraries with Radiolabeled Acceptor Molecules. *Pept. Res.* 1993, *6*, 129–133.
(152) Lebl, M.; Patek, M.; Kocis, P.; Krchnak, V.; Hruby, V. J.; Salmon, S. E.; Lam, K. S. Multiple Release of Equimolar Amounts of Peptides from a Polymeric Carrier Using Orthogonal Linkage-Cleavage Chemistry. *Int. J. Pept. Protein Res.* 1993, *41*, 201–203.
(153) Dower, W. J.; Barrett, R. W.; Gallop, M. A. Method of Synthesizing Diverse Collection of Oligomers. PCT Application WO 93/06121.
(154) Needels, M. N.; Jones, D. G.; Tate, E. H.; Heinkel, G. L.; Kochersperger, L. M.; Dower, W. J.; Barrett, R. W.; Gallop, M. A. Generation and Screening of an Oligonucleotide-Encoded Synthetic Peptide Library. *Proc. Natl. Acad. Sci. U.S.A.* 1993, *90*, 10700–10704.
(155) Lerner, R.; Janda, K.; Brenner, S.; Nielsen, J. Encoded Combinatorial Chemical Libraries. PCT Application WO 93/20242.
(156) Brenner, S.; Lerner, R. A. Encoded Combinatorial Chemistry. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 5181–5183.
(157) Nielsen, J.; Brenner, S.; Janda, K. D. Synthetic Methods for the Implementation of Encoded Combinatorial Chemistry. *J. Am. Chem. Soc.* 1993, *115*, 9812–9813.
(158) Kerr, J. M.; Banville, S. C.; Zuckermann, R. N. Encoded Combinatorial Peptide Libraries Containing Non-Natural Amino Acids. *J. Am. Chem. Soc.* 1993, *115*, 2529–2531.
(159) Nikolaiev, V.; Stierandova, A.; Krchnak, V.; Seligmann, B.; Lam, K. S.; Salmon, S. E.; Lebl, M. Peptide-Encoding for Structure Determination of Nonsequenceable Polymers Within Libraries Synthesized and Tested on Solid-Phase Supports. *Pept. Res.* 1993, *6*, 161–170.
(160) Ohlmeyer, M. H. J.; Swanson, R. N.; Dillard, L. W.; Reader, J. C.; Asouline, G.; Kobayashi, R.; Wigler, M.; Still, W. C. Complex Synthetic Chemical Libraries Indexed with Molecular Tags. *Proc. Natl. Acad. Sci. U.S.A.* 1993, *90*, 10922–10926.
(161) Borchardt, A.; Still, W. C. Synthetic Receptor Binding Elucidated with an Encoded Combinatorial Library. *J. Am. Chem. Soc.* 1994, *116*, 373–374.

JOURNAL OF
MEDICINAL
CHEMISTRY

Volume 37, Number 10 May 13, 1994

Perspective

Applications of Combinatorial Technologies to Drug Discovery. 2. Combinatorial Organic Synthesis, Library Screening Strategies, and Future Directions[1]

Eric M. Gordon,* Ronald W. Barrett, William J. Dower, Stephen P. A. Fodor, and Mark A. Gallop

Affymax Research Institute, 4001 Miranda Avenue, Palo Alto, California 94304

Received March 1, 1994

Biographies

Ronald W. Barrett received his Ph.D. in Pharmacology from Rutgers University in 1983 and did postdoctoral work at the Addiction Research Foundation in Palo Alto, CA, in the field of opiate receptors. From 1986 to 1989, he worked in the Neuroscience Research Area at Abbott Laboratories in Abbott Park, IL. He joined Affymax in 1989 where he established the Receptor Pharmacology Group. Dr. Barrett is presently Vice President and Director of Receptor Pharmacology.

William J. Dower received his Ph.D. in Biology from the University of California, San Diego, and did postdoctoral work at Stanford University studying steroid control of specific mRNA metabolism and the amplification of genes mediated by chemotherapeutic agents. He joined Bio-Rad Laboratories in 1984 where as a member of the newly formed molecular biology unit, he introduced the electroporation technique for the high-efficiency transformation of bacteria. He joined Affymax in 1989 to establish a molecular biology research group. Dr. Dower is Director of Molecular Biology at Affymax.

Stephen P. A. Fodor received his Ph.D. in Chemistry at Princeton University and was an NIH postdoctoral fellow in Chemistry at the University of California, Berkeley. He joined Affymax in 1989 as a Staff Scientist in Optical Technologies, becoming Director of Physical Sciences. Dr. Fodor's group led the development of new technologies, merging photolithography with combinatorial solid-phase chemistry. In 1993 he joined Affymetrix as Scientific Director, where he is currently using oligonucleotide arrays to study a variety of DNA molecular recognition processes.

Mark A. Gallop received his Ph.D. from the University of Cambridge, England, for research in the area of organo transition metal chemistry and did postdoctoral work in the laboratories of Peter G. Schultz at the University of California, Berkeley. He joined Affymax in 1990 and is currently a Senior Scientist and leader of the Combinatorial Chemistry Group.

Eric M. Gordon received his Ph.D. in 1973 from the University of Wisconsin—Madison and did postdoctoral work at Yale University. His research interests include the rational design of enzyme inhibitors, the chemistry of amino acids, peptides, and natural products, and combinatorial chemistry. Dr. Gordon has authored 150 papers and U.S. patents in these areas. From 1974 to 1992 he worked at the Squibb Institute for Medical Research (subsequently the Bristol-Myers Squibb Pharmaceutical Research Institute) in Princeton, NJ, most recently as a Director of Medicinal Chemistry. In 1992, he joined the Affymax Research Institute where he is currently Vice President of Research and Director of Chemistry.

A. Combinatorial Organic Synthesis

The notion of creating huge, searchable libraries of small organic molecules is unprecedented in medicinal chemistry, and the possibility of doing so has recently captured the imagination of the drug-discovery community. The conventional paradigm of small molecule lead development, in which a compound undergoes many rounds of individualized, hand-crafted modification and biological testing en route to drug candidacy, will likely be dramatically accelerated by the application of combinatorial chemistry technologies to mass-produce and evaluate lead analogs. The ability to harness molecular diversity techniques as tools for lead discovery offers an unparalleled opportunity for medicinal chemistry to expand the breadth and scope of molecular structures that may be screened for biological activity. Widespread availability of collections of highly diverse small-molecule libraries should provide an opportunity to assess the impact of combinatorial organic synthesis on new-lead discovery. In this section of part 2, some of the issues which confront the practitioner of combinatorial organic synthesis, as they relate to the problems of molecular recognition in general and medicinal chemistry in particular, will be analyzed.

Issues in Practicing Combinatorial Organic Synthesis

Combinatorial organic synthesis (COS) presents somewhat of an intellectual inversion of the past 50 years of synthetic organic chemistry. The chemist-of-the-Woodwardian era was interested in a masterful and carefully

0022-2623/94/1837-1385$04.50/0

plotted natural product synthesis of a complex entity of known structure. Reactions were more often employed or developed to solve specific challenges rather than to provide generic methodologies. Rigorous control of reaction pathways, stereochemistry, and regiochemistry, and the exclusion of all but the desired diastereomer were obligatory in a faithful rendering of the technique. In contradistinction to natural product total synthesis, rather than generating a single, specific entity, the goals of COS are to create *populations* of molecular structures. Rather than exercising complete control, the combinatorial chemist, while maintaining high reaction efficiency and relative reactive compatibility, may actually seek to create situations and apply strategies in which stereochemical/regiochemical control is relaxed. This must be achieved while remaining cognizant of the impact these factors may have on the stoichiometry of the resulting libary and its design and ultimate use. Hence, the combinatorial chemist seeks to apply a series of Woodwardian reactions (reliable, high yielding) that operate generically on a diverse set of building blocks to provide a multitude of related products.

Criteria for Library Design. The primary objectives of producing small-molecule libraries by COS are to provide collections of compounds suitable for both drug-discovery screening and drug-development optimization. When complete, the combinatorial drug-discovery exercise should have created a stable population of low molecular weight entities, free of reactive and toxicity-causing functionality. While a paramount medicinal chemistry design criteria for small-molecule-libary construction might be that the *products* of diversity generation (individual library members) should "look" like drug leads, of still greater importance is that the library actually contains compounds capable of interacting at some detectable level with the biological target of interest. When small-molecule leads for a target have been previously defined (e.g., benzodiazepine ligands for a peptide or other G-protein-coupled receptor, transition-state inhibitors for a protease), the notion of searching for more potent derivatives among libraries combinatorially enriched in specific pharmacophore analogs is an obvious tactic to pursue. However, as the universe of well-defined macromolecular drug targets continues to expand through the impact of molecular cloning, the problem of identifying new pharmacophores capable of modulating the various interactions of peptides, proteins, carbohydrates, oligonucleotides, or lipids at these sites will also be intensified.

Will "rules" about the types of libraries that may prove generally useful in ligand discovery be discovered? Although the field of molecular diversity has not yet matured to the point where substantial insight into this question is forthcoming, it is intuitively obvious that small-molecule libraries, whose members structurally resemble historical leads, should provide a fertile reservoir of potential molecular diversity. Tangential to this, natural products aside, numerous historical drug leads were derived simply because synthetic routes to these molecules were readily available. It is likely that early-stage COS will be limited by applicable chemistry and that this will necessarily focus work toward traditional leads, whose syntheses are known and well-documented.

The successful identification from recombinant libraries of L-amino acid-based peptide ligands that inhibit protein–peptide, protein–protein, and protein–carbohydrate interactions suggests there is broad utility in screening large libraries of peptidic compounds. It remains to be seen whether collections of other random molecular structures that are quantitatively as diverse as existing peptide libraries prove in *de novo* ligand discovery to include the "pharmacophores of the future".

Ligand rigidity may be another important parameter to consider in the course of library design. The incorporation of conformational constraints into flexible lead molecules has emerged as a powerful strategy to enhance ligand potency and/or selectivity, particularly in the field of peptidomimetic medicinal chemistry.[2-10] Nevertheless, with regard to library design, conformational restriction may act as a two-edged sword: an inappropriate constraint is likely to abrogate the modest but perhaps detectable activity of a more flexible analog, which could, in a secondary library, be systematically constrained. From the point of view of random screening, it remains to be determined whether useful leads will arise more frequently from libraries of rigidified or flexible structures. Data from the evaluation of cyclic peptide libraries in both synthetic and recombinant systems may provide some important insights into this issue. A number of methods have been recently described for on-resin cyclization of peptides through both main-chain and side-chain functional groups.[6-10,68] At present, a portfolio of libraries containing both conformationally rigid and relaxed molecular diversity seems most appropriate. A longer range solution might be to moderate the high risk of conformational restriction by creating very large populations of semirigid molecular arrays, comprising structural families that collectively sample as completely as possible all regions of conformational space.

Characterization. The usual measures of evaluating success in organic synthesis may lose meaning in COS. The classical notions of such fundamental concepts as purity/homogeneity, yield, exact product structure, relative and absolute stereochemical control, specific physical properties are less relevant when applied to a broad population of molecules (of course they may become quite relevant as individuals emerge from a selection process). Additionally, the analytical mainstays of the synthetic organic chemist, such as NMR and IR, may become obviated. The NMR spectrum of a 10 000-component library mixture is not diagnostic. The loss of these powerful tools requires that compensating technologies be developed. A major dilemma of COS is the difficulty of confirming the degree to which the expected chemistry has proceeded on the entire population of substrate molecules. Several groups have recently reported on the use of electrospray mass spectrometry as a technique for evaluating the bulk composition of diverse peptide libaries.[11,12] Gross synthetic discrepancies, such as incomplete protecting group removal, may be detectable by mass analysis, providing an opportunity to optimize the library synthesis protocols. In the characterization of combinatorial products, the presence of "byproducts" (in COS, unexpected products), combined with the difficulty of detecting these compounds, will cause problems if one mistakenly concludes that a screening hit is the expected product. This section will conclude by offering a potential solution to this problem.

Efficiency/Automation. Among the chemical criteria relevant for small-molecule-library design is the efficiency of diversity creation. The assembly of most small molecules reduces to the intercombination of only three to

five building blocks of molecular weight ~150 each. Synthetic reactions capable of combining numerous building blocks simultaneously constitute a highly efficient form of diversity generation. Thus the Ugi four-component reaction has a high combinatorial efficiency since building blocks of four families (amines, carbonyl compounds, isocyanides, and suitable acid components) are linked simultaneously to afford α-amino acid derivatives. In contrast, peptide chemistry traditionally links two building blocks at a time. In both the broad screening and the lead analoging modes, a longer range question pertains to the ability of the chemistry to eventually be automated. Once the key decisions and overall strategy have been determined, much of the actual chemistry is repetitive in nature. Machines will continue to be constructed to capitalize on this and libraries will be assembled under computer control.[13-15]

Quantity and Quality of Diversity. While the "quantity of diversity" that is experimentally accessible can be dictated by the number of building blocks in the basis set and by the number of synthetic operations applied, or able to be applied (see part 1[1]), the practical limitations on library size are most generally imposed by the format within which the diversity is created and evaluated. A small number of building blocks subjected to many synthetic steps will yield high (numerical) diversity; however the products of these reactions may be relatively large molecules, not well-suited for lead development as traditionally administered therapeutics. Thus, as the combinatorial process proceeds, an opportunity window may exist in which the bulk of the library possesses properties which standard medicinal chemistry usually seeks in small-molecule drug discovery (MW < 700, solubility, etc.). Continued application of the combinatorial process will lead to product libraries containing larger molecules (composed of more building blocks) wherein the individual library members have "outgrown" the classical criteria of a lead-drug molecule.

In surveying the historical landscape of drug discovery, there are particular pharmacophores or structural arrays which periodically surface far in excess of random chance (benzodiazepines, β-lactams, imidazoles, phenethylamines, etc.).† A review of recent successes in the era of "rational drug design" suggests that certain molecular concatenations—protein turn mimetics, conformationally restricted amino acids, transition-state analogs, dipeptide isosteres, molecular scaffolds, designed elements for enzyme inhibition—are often found in the medicinal chemistry of lead compound development. In consideration of the molecular structures which have left their mark on modern medicinal chemistry, one might conclude that the drug-discovery process is impacted not only by the sheer quantity of diversity surveyed, but additionally by the more subjective "quality" of diversity that is evaluated. Different organizations and individuals will certainly bring a wide variety of criteria to the subjective appraisal process, depending on style, experience, and bias. It may be speculated that the quality of diversity will be influenced by the sophistication of the building blocks originally introduced into the combining system (library bias on the part of the medicinal chemist) and the extent to which molecular substructures of the building blocks can be assembled in diverse, spatial (3-D) relationships. Thus the collected expertise of medicinal chemical knowledge may be used to "hyperevolve" or "bias" the library by the planned introduction of commonly evolved elements; these elements are "retrocombinatorial synthons" of many known bioactive classes. Thus the building block basis set must be judiciously chosen and carefully attuned to the collected knowledge historically amassed in drug discovery.

† An interesting aside regarding these important substructures is that development of "generic" syntheses of key pharmacophores ultimately enabled facile generation of many analogs. Concurrently or subsequently, diverse biological activities were found among these compound classes. In a sense, this is suggestive of combinatorial chemistry, except the crucial molecules were made serially rather than in a parallel/combinatorial high throughput fashion.

Issues in the Selection of Building-Block Sets

The acquisition of a building-block library can be a major time and resource investment, and the eventual decision of which type of chemical building blocks to utilize places limits on the universe of structural diversity which ultimately can be explored. Depending on the specific objective, important building-block criteria include the availability of a large number of diverse, fairly complex, easily accessible starting materials. These may be either commercially available or prepared in a few steps from commercial materials. Members of a building block set should reflect a broad array of physicochemical properties, functionality, charge, conformation, etc. Building blocks may be chiral, achiral, or racemic. Certain building-block families have what may be termed a high "combinatorial potential". This relates to the high density per carbon atom of reactive functionality which can participate in new covalent combinations. For example, monosaccharides have high combinatorial potential since the high density of available hydroxyl groups leads to many potential connecting permutations. In addition to polymer formation, the high combinatorial potential of such types of building blocks may also be exploited as scaffolds for the generation of diversity (*vide infra*).

Synthetic Strategy

An important strategic element in combinatorial library synthesis is the degree of reliability of the ligand synthesis chemistry. What is the likelihood of general synthetic success with a particular reaction? The nature of combinatorial reactions, which must proceed in the face of a broad range of functionality on a multitude of substrates and where the products are difficult to analyze individually, demands that, in selection of synthetic methodologies, greater weight must be given to reaction sequences with reliable, predictable outcomes. A more subtle question revolves around the number of synthetic options available in the course of diversity generation. For example, a synthetic strategy structured in such a way that, as the process proceeded, new combinatorial possibilities opened up, would be preferable to having options narrow, especially if the goal was generating a maximum of structural diversity.

As previously noted, there are two distinct themes that must be considered for the successful application of combinatorial technologies to ligand discovery and optimization, *viz.* broad-based screening and directed chemical analoging. The issues underlying conceptual design, as well as the synthetic strategies utilized in construction of these different classes of libraries, are noteworthy and are summarized in Figure 1. Building block requirements for undertaking broad and narrow diversity searches differ markedly. The search for an initial lead molecule may be essentially a random screening exercise, where the em-

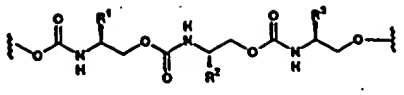
Broad Screening	Chemical Analoging/Optimization
huge size library	modest size library
broadest structural diversity	relatively narrow structural diversity
no special initial structure goal	specific structural goal
any building blocks	specific retrocombinatorial building blocks
undefined order of reaction	specific order of combination
flexible synthetic strategy	well defined synthetic strategy
site of tether not crucial	tether crucial-build in redundancy
ligand possibly uncouplable	ligand should be releasable
single selection evolution	cumulative selection evolution

Figure 1. Combinatorial chemistry: comparison of two major themes.

phasis is on exposing a macromolecular drug target to the maximum possible structural diversity. The objective is to identify a ligand of significant affinity for the target, the exact ligand structure and its detailed characteristics at this point are not relevant: in fact any molecule will do. An approach to generating highly diverse libraries for use in medicinal chemistry might favor using building blocks which have distinguished themselves by appearing frequently in previous active leads (e.g., statine, hydroxyethylamines, Freidinger lactams[16]). On the other hand, once a lead is available, most drug discovery proceeds through a series of evolutions (optimizations) in order to meet a set of predetermined criteria. Since specific structural types are sought, searching in a very broad pool of diversity (as above) is unlikely to be successful (actually it could uncover a new lead but is less likely to optimize an existing one). Ideally, what is required in this type of diversity-generating strategy is to "explode" around the known lead, i.e., to create as highly diverse population as possible that bears close structural resemblance to the original hit, followed by a selection for desired criteria.

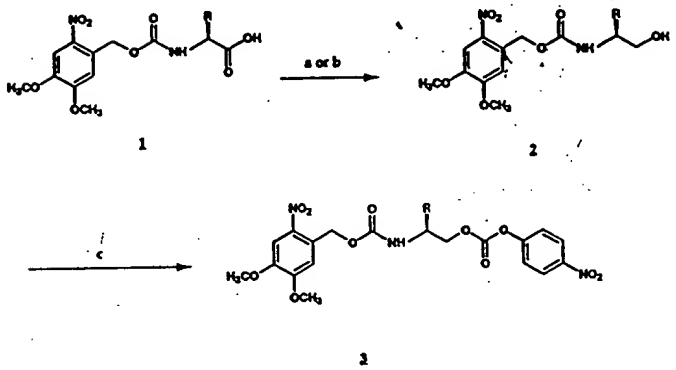
Clearly the subunits which lead to predetermined structures must be quite specific: from where should building blocks for known structural classes of pharmacophores arise? The answer, as in organic synthesis, lies in a retrosynthetic analysis or what we may term a *retrocombinatorial* approach to building-block selection. Lead structures should be retrosynthetically dissected in the maximum number of ways and upon these various possibilities imposed the needs of performing combinatorial chemistry. Inspection of the retrosynthetic tree invites the following key questions: By which modes of forward synthesis are the most building blocks available or obtainable? If the synthesis is allowed to proceed by that course, what is the scope and degree of reliability of the necessary reactions? Extending this line of reasoning should permit the maximum leverage to be applied combinatorially.

A common feature of both paradigms is likely to be a reliance on solid-phase-synthesis methods to facilitate the assembly of combinatorial libraries. Synthesis on a polymeric support greatly simplifies the problem of product isolation from reaction mixtures and also facilitates the partitioning of products into multiple aliquots for subsequent chemical elaboration. Moreover, the opportunity exists to take advantage of the support-tethered diversity in the design of convenient receptor-binding assays for library evaluation. While there has been a long tradition of polymer-supported organic chemistry,[17-20] it is only in the areas of peptide and



Polycarbamates

Figure 2. Structure of a synthetic oligocarbamate prototype.



3

Figure 3. Synthesis of N-protected nitrophenyl carbonate monomers. Key: (a) BH_3, THF; (b) DCC, methylene chloride, *N*-hydroxysuccinimide, HOBt; then sodium borohydride, ethanol; (c) *p*-nitrophenyl chloroformate, methylene chloride, pyridine.

oligonucleotide synthesis that solid-supported chemistry has truly been optimized and become common-place. The advent of combinatorial organic synthesis will undoubtedly signal a renaissance in solid-phase organic chemistry, as workers attempt to adapt well-characterized homogeneous reactions to reliable solid-supported protocols.

Progress to Date: Synthetic Polymeric Diversity

The design and synthesis of novel synthetic monomers which, when assembled in a combinatorial fashion, could yield relatively low molecular weight polymeric materials is an approach that is well-suited to diversity generation and evaluation. Combinations of such monomers could lead to substances with novel backbones, possibly possessing desirable properties, such as metabolic stability, enhanced pharmacokinetic profiles, and cell and membrane permeability. Identification of these and other potentially modifiable parameters in such systems could facilitate drug discovery.

Schultz and co-workers have reported the synthesis of a library of oligocarbamates starting from a basis set of chiral aminocarbonates[21] (Figure 2). The monomeric units were readily obtained by the modification of amino acids via the intermediacy of the corresponding chiral amino alcohols (see Figure 3). The resulting nitrophenyl carbonate building blocks (*3*) were stable for several months at room temperature.

Oligocarbamates were synthesized on a solid support by deprotection of a resin-bound amine, protected with either the base-labile Fmoc or photolabile nitroveratryloxycarbonyl (Nvoc) group, followed by treatment with a nitrophenyl carbonate of type 3. The deprotection/coupling cycle was repeated until an oligocarbamate of the desired length was attained (seven or eight cycles). Overall coupling yields were greater than 99% per step. Side-chain deprotection followed by resin cleavage afforded the desired oligocarbamates (Figure 4).

The VLSIPS photolithographic chip format, previously employed for oligopeptide synthesis, was used in the construction a spatially-addressable oligocarbamate library of 256 members. An anti-carbamate monoclonal antibody served as a model receptor for screening against this array. Antibody:oligocarbamate complexes were

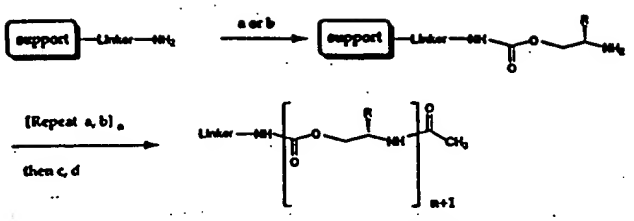


Figure 4. Solid-phase synthesis of oligocarbamates. Key: (a) nitrophenyl carbonate monomer, HOBt, diisopropylethylamine, NMP; (b) piperidine, NMP or *hν*; (c) acetic anhydride, NMP; (d) TFA, triethylsilane.

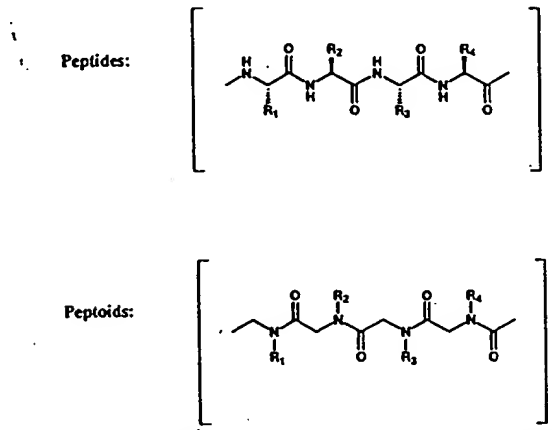


Figure 5. Comparison of peptide and peptoid backbones.

detected by treatment of the chip with a fluorescein-conjugated secondary antibody, followed by analysis using scanning epifluorescence microscopy. Because the location and structure of each different library member is defined by the synthetic strategy (binary masking) used in this technique, the necessity of sequencing the products is obviated. The binding activities of putative hits were confirmed by conventional assays using authentic material prepared by independent synthesis. A preliminary evaluation of the physicochemical properties of oligocarbamate molecules indicate that they are more hydrophobic than the corresponding peptide homologs, and their expected resistance to several proteases was confirmed.

Another type of synthetic polymeric diversity has been developed by Simon *et al.*[22,23] Through a variety of preparative routes, this group created a basis set of monomeric N-substituted glycine units, each bearing a nitrogen substituent similar to those of the natural α-amino acid side chains. The formal polymerization of these monomers results in a class of polymeric diversity which these workers have termed "peptoids" (Figure 5). Peptoids may be synthesized either manually or robotically following either a "full monomer" oligomer synthesis or via a "submonomer" synthesis, as reported by Zuckermann *et al.*[24] and illustrated in Figure 6. Various biological activities have been established for specific peptoid sequences, including inhibition of α-amylase and the hepatitis A virus 3C protease, binding to the tat RNA of HIV[22], and antagonism at the α_1-adrenergic receptor.[25] The peptoid approach to diversity generation has been extended to the preparation of encoded combinatorial libraries, in which natural amino acids code for the structure of the peptoid chain[26] (see part 1[1] and Figure 7).

An important variant of the synthetic polymeric diversity approach is directed toward construction of a chemical library in which the peptidyl backbone is conserved but a dipeptide unit is replaced at specific

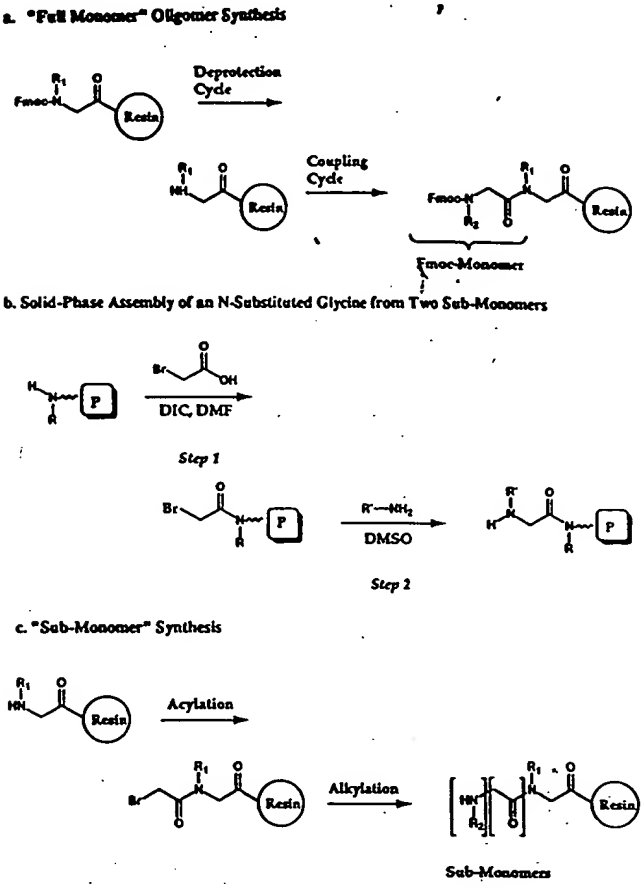


Figure 6. Synthesis of peptoids.

positions by a phosphonate dipeptide surrogate (see Figure 8). Such phosphonate pharmacophores are well-known as transition-state analogs for amide bond cleavage and have found wide usage in the inhibition of metalloproteases.[27-29] Campbell has described methods for the solid-phase synthesis of peptidylphosphonates that are compatible with the Fmoc/^{t}Bu protecting group strategy of standard peptide synthesis.[30] The key reaction step is formation of the phosphonate ester bond, which is achieved via a modified Mitsunobu condensation (Figure 9). Precursor lactic acid and protected amino phosphonate building blocks are prepared as shown in Figure 10.

When this process is applied to the combinatorial synthesis of peptidylphosphonates, the diversity product will be a metalloprotease enzyme inhibitor library. Enzyme-inhibitor libraries of this type and those focusing on other known inhibitory pharmacophores (thiols, hydroxamates, carboxyalkyldipeptides, etc.) may prove to be important tools in rapidly profiling novel proteases and for determining which pharmacophores are most effective at their inhibition. Using this knowledge, secondary inhibitory libraries can be constructed to optimize original leads. Through such a process it may be possible to dramatically accelerate the process of finding high-affinity enzyme-inhibitor ligands.

Another interesting type of polymeric diversity based upon a vinylogous polypeptide backbone has recently been reported by Hagihara *et al.*,[31] in which introduction of a trans olefinic linkage between the α-carbon and the carbonyl group of various amino acids is generalized. Additionally, Smith and colleagues have synthesized a non-amide polymer of (3,5)-linked pyrrolin-4-one oligomers which mimic the β-strand conformation of a normal peptide chain[32] (see Figure 11).

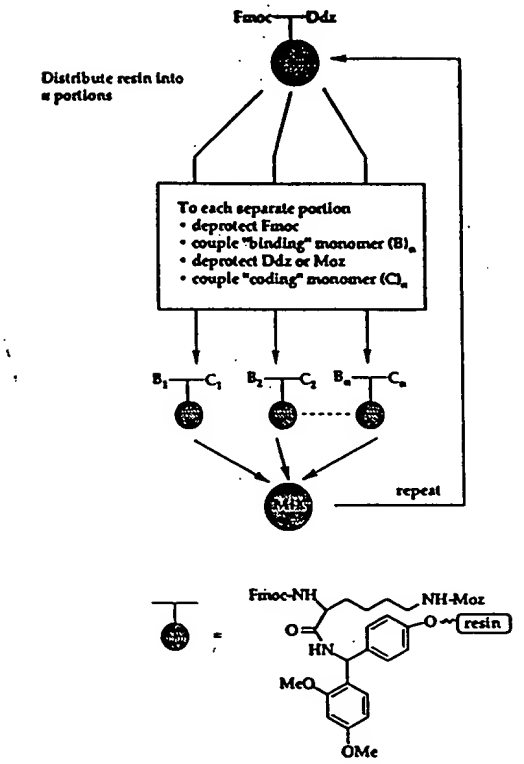


Figure 7. The synthesis of an encoded library consists of the following steps: (1) A bifunctional linker containing two orthogonally protected points of extension (Nα-Fmoc-Nε-Moz-Lys-OH) is attached to polystyrene resin via an acid-labile linker. (2) The solid support is divided into n equal portions at a mixture position. (3) A unique Nα-Fmoc-protected non-natural monomer (B) is coupled to the "binding" strand. (4) A series of Nα-Ddz-protected L-amino acids (C) are then coupled to the "coding" strand. (5) The solid supports are recombined.

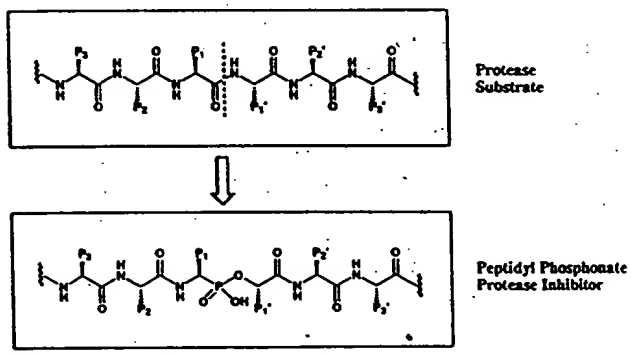


Figure 8. Peptidyl phosphonates.

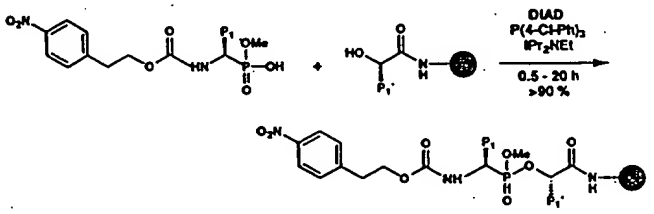


Figure 9. Solid-phase peptidyl phosphonate synthesis (SPPPS).

Nonpolymeric, Small-Molecule Diversity. The majority of chemical diversity generation discussed above concerns the preparation of linear molecules, in which the target structures are unambiguously specified by the order of building-block addition. In contrast, the great preponderance of organic synthesis proceeds rather differently, wherein building blocks interlock to give rise to

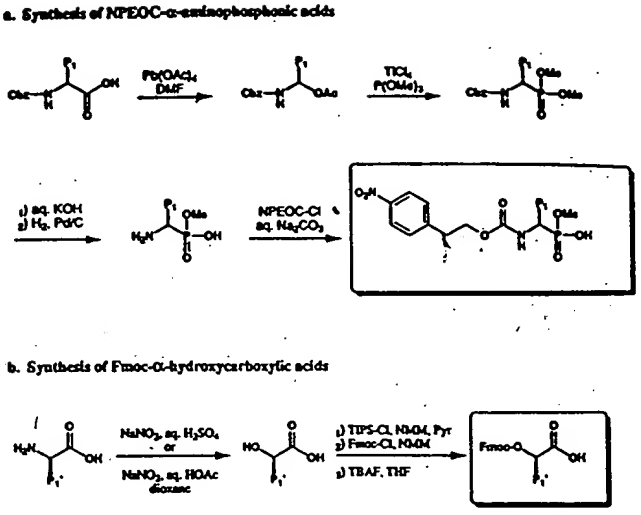


Figure 10. Synthesis of peptidyl phosphonate building blocks.

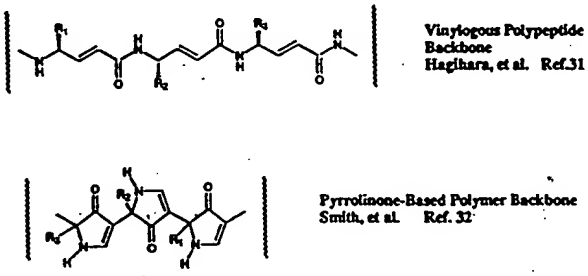


Figure 11. Novel polymeric backbones.

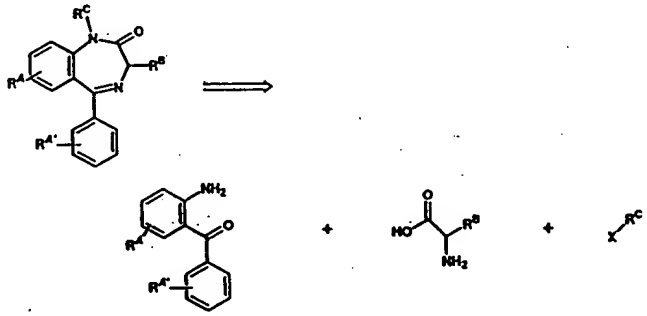


Figure 12. Components of a benzodiazepine library.

nonpolymeric, three-dimensional arrays. The recent seminal work of Ellman on the solid-phase synthesis of 1,4-benzodiazepines lays the groundwork for creation of a small-molecule library of one of medicinal chemistry's most notable pharmacophores and represents one of the first examples of the application of combinatorial organic synthesis to nonpolymeric organic compounds.[33]

The benzodiazepines are synthesized on a solid support by the connection of three building blocks, each of different chemical families (Figure 12). Following the attachment of 2-aminobenzophenone hydroxy or carboxy derivatives to the support using an acid-cleavable linker, [(4-hydroxymethyl)phenoxyacetic acid], the N-protecting group is deblocked (piperidine/DMF), and the weakly nucleophilic amine is acylated with an α-Fmoc-protected amino acid fluoride, using 4-methyl-2,6-di-*tert*-butylpyridine as an acid scavenger (Figure 13). Fmoc deprotection, followed by treatment with 5% acetic acid in DMF, causes the general cyclization to the intermediate lactam. Capitalizing on the ability of lithiated 5-(phenylmethyl)-2-oxazolidinone to selectively deprotonate the anilide NH,

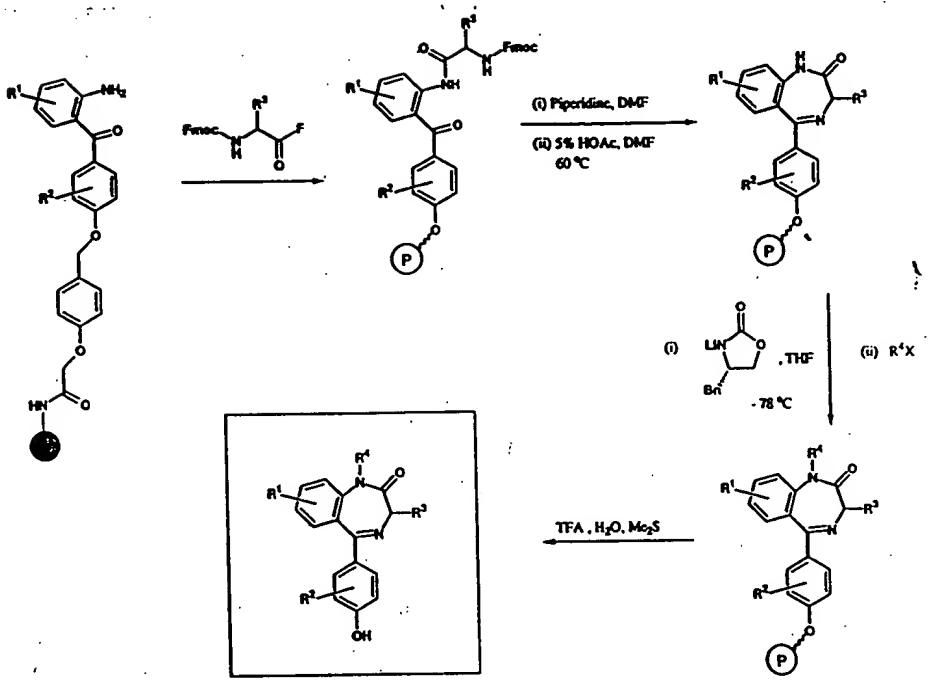


Figure 13. Combinatorial synthesis of the benzodiazepine pharmacophore.

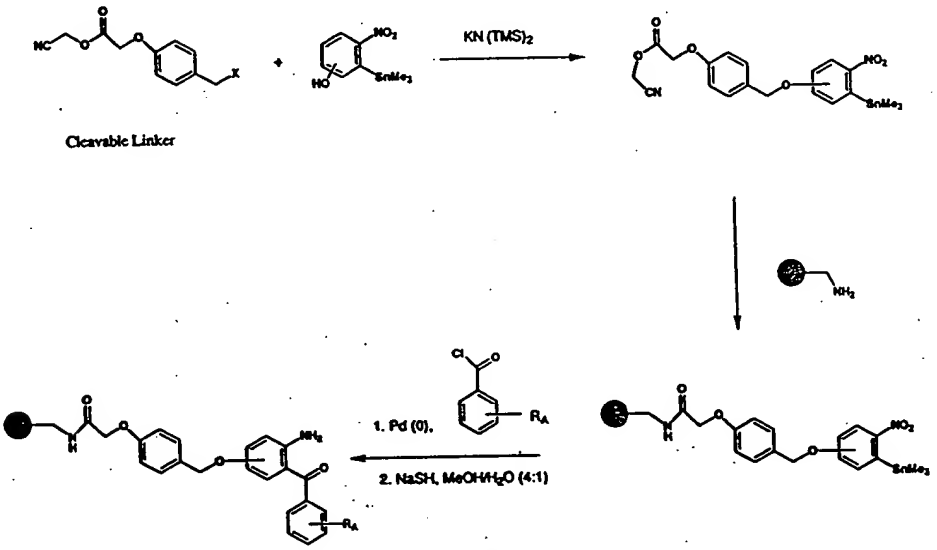


Figure 14. Synthesis of 2-aminobenzophenone derivatives on a solid support.

alkylation was achieved with a variety of alkylating agents. Aqueous acid cleaves the new benzodiazepine from the support in very high overall yields. The integrity of the chiral center was confirmed by a racemization test using chiral HPLC.

One of the limiting features of applying the above scheme to combinatorial library construction is that, though many alkylating agents and amino acid building blocks are commercially available, there is not a ready supply of appropriately functionalized 2-aminobenzophenones. Ellman addressed this problem directly by creating a general method for preparation of these materials on solid supports[34] (Figure 14). The stage is now set for the Ellman laboratory to create a benzodiazepine library.

Several other approaches to nonpolymeric molecular diversity have recently been published. In pursuit of small-molecule libraries, Nikolaiev *et al.* have used their amino acid encoding format (part 1[1]) with a building block basis set combining both amino acids and other synthetic units to prepare collections of nonpeptidic compounds and peptides refractory to Edman degradation (N-blocked peptides).[35] Representative examples of molecules which have emerged from such non-peptide libraries are shown in Figure 15.

A feature of several of the formats used in the display of synthetic diversity is that the potential ligands are tethered to a solid support. While screening strategies have been developed to exploit this feature, it is frequently desirable to screen compounds in solution. Many groups have engaged in developing releasable linker strategies to solubilize potential ligands. The issue has been addressed by a considerably different strategy by Hobbs DeWitt *et al.*, in which solid-phase chemistry, organic synthesis, and a designed parallel reaction apparatus were utilized for the generation of small-molecule libraries, the individuals of which, were termed "diversomers".[36] Target compounds which included dipeptides, hydantoins, and benzodiazepines were synthesized simultaneously but separately,

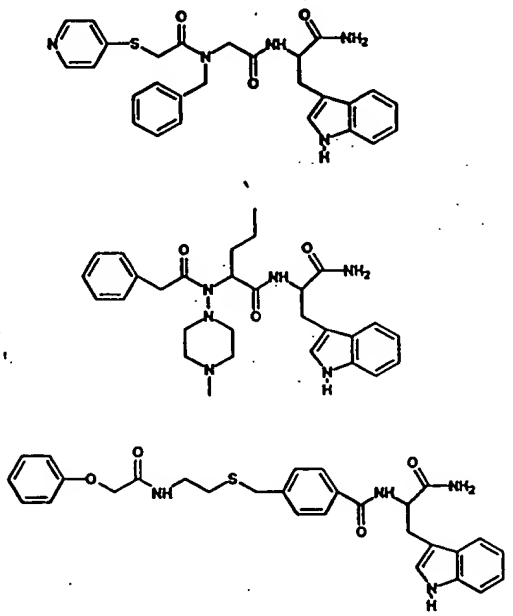


Figure 15. Structure of representative molecules from the Nikolaiev *et al.* nonpeptide library (ref 35).

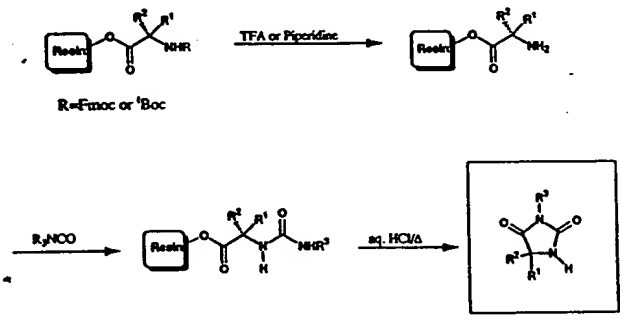


Figure 16. Synthesis of a hydantoin library.

on a solid support in an array format, to generate a collection of up to 40 discrete structurally related compounds. The preparation of hydantoins was carried out as shown in Figure 16. The synthetic strategy is directed through a resin-tethered penultimate product, in which the orchestrated revealing of distal functionality initiates attack on the resin-linking bond to eject the newly formed product into solution. Products which fail to react, should remain attached to the solid phase, and thus aid in product purification. The yields of hydantoins released from the resins in the final step ranged from 4 to 81% on a scale of 0.3–11 mg, which should be sufficient to support most preliminary *in vitro* biological testing. The resulting soluble, small molecules were characterized by traditional means. The authors also note the utility of ^{13}C gel-phase NMR to monitor reaction progress of the resin-bound intermediates.[37,38]

In a similar manner, a general method for multiple, simultaneous synthesis of soluble benzodiazepines was developed (Figure 17). Eight groups of five-amino acid resins were trans-imidated with five groups of eight 2-aminobenzophenone imines to form 40 resin-bound imines. Treatment with TFA liberated 40 discrete benzodiazepines from the resins. The products were obtained in 2–14-mg quantities, corresponding to 9–63% yields with estimated purities of >90%. Though the numbers of compounds involved in the diversomer methodology (~40) are significantly smaller than that which can be prepared by other library strategies (10^4–10^8), this interesting

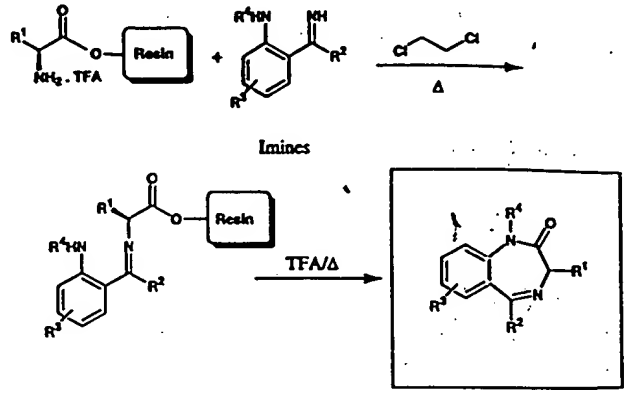


Figure 17. Preparation of a soluble benzodiazepine library.

approach to parallel organic synthesis produces relatively pure materials on a preparative scale in the traditional, soluble format.

Future Innovations

As the field of combinatorial chemistry receives increasing attention from the pharmaceutical establishment, it seems likely that the contents of chemical libraries will continue to evolve to look more and more like the type of compounds which have previously led to drugs. In spite of the complexity which early parts of the process may pose to the combinatorial chemist, a hidden advantage which combinatorially-derived molecules offer is that any "hit" will be readily synthesizable, by definition. This should be contrasted with a natural product driven approach to drug discovery and development, where often the structural complexity of the lead compound hampers the rapid preparation of analog molecules and the acquisition of SAR.

A previous point deserving further emphasis is that the vast universe of synthetic organic reactions are idiosyncratic transformations that fail to afford quantitative yields of unique products. Most synthetic chemistry procedures afford multiple products (regio- and stereoisomers, etc.) in variable yields. If diversity-generating chemistry proceeds ambiguously, how then are the results of small-molecule combinatorial organic syntheses to be understood and appropriate information extracted from library analysis? It may be speculated that encoding techniques will provide one method by which the combinatorial organic chemist can address the practical inefficiencies of chemical synthesis. Instead of envisioning an encoding tag as explicitly specifying the structure of an associated entity, one might consider the tag as *a record of the chemical history of individual library members*. Thus, after encoding the "recipe" or synthetic protocol used in the assembly of a combinatorial library, the library may be screened for active recipes. Once identified as "active", the synthesis would be replicated on a preparative scale, and the product mixture fractionated to identify active product(s). This strategy shifts emphasis from the criterion of singularity in a reaction outcome (a single predictable structure) to reproducibility and compatibility (orthogonality) with chemistry used in the synthesis of the encoding tag and in preparative scaling. The creation of encoded, small-molecule diversity, which can be released from a support (solubilized) while some type of link to the original tag is also maintained, is also likely to be an important area of investigation.

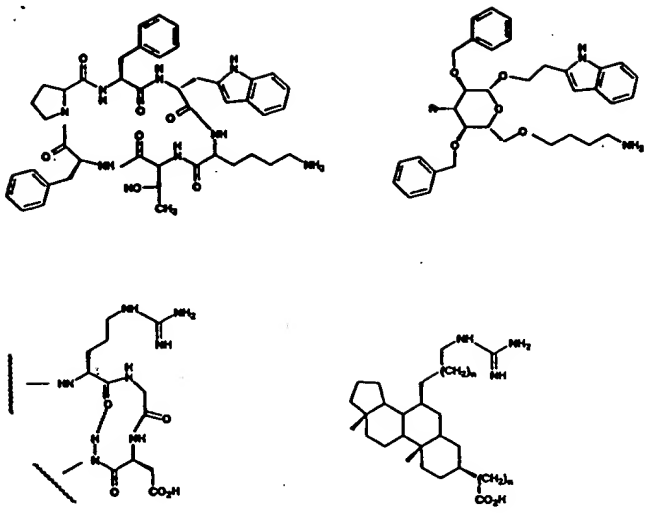


Figure 18. Scaffolds as templates for combinatorial chemistry.

One type of noteworthy chemical strategy which may have a bright future in the combinatorial realm rests on the conceptual extension of the work of Hirschmann, Nicolaou, and Smith and their co-workers into designing nonpeptidal peptidomimetics by the appropriate functionalization of designed scaffoldings (Figure 18). A specific example of this approach is the design and synthesis of a β-D-glucose-based nonpeptide mimetic of a potent cyclic hexapeptide somatostatin agonist.[39,40]

Hirschmann and co-workers have also used a functionalized steroidal template to serve as a backbone for mimicking a type 1 β turn.[41,42] The fact that such a remarkable job of molecular mimicry can be achieved with individual compounds bodes well for the application of this approach to combinatorial methods.

No library will ever be "complete" but instead will sample a subsection of a particular universe of molecular structure and space. In certain situations, libraries may be considered to be starting materials for the construction of new libraries of diversity. It is also useful to consider chemical libraries as collectable or archivable entities. Ideally, one seeks to preserve new compound libraries and use them for a variety of present and future screening needs. As time passes, the combinatorial chemist will be in possession of an accumulating collection of molecular diversity with which to challenge new drug targets. If chemical libraries are to become an item of commerce, a good deal more will have to be learned about their "shelf life" and how best to store them for future use. To date, the shelf life of chemical libraries is an open question.

B. Methods for Screening Combinatorial Libraries

The importance of distinguishing between the two principal applications of combinatorial technologies in ligand discovery, i.e., broad screening versus directed analoging, is particularly relevant to the design of assay methodologies for library evaluation. In searching large, highly diverse libraries for novel lead compounds, a premium is placed on the ability to detect rare ligands that may have modest affinity for the target receptor. The assay strategy may differ in screening analog libraries, since one is trying to develop quantitative SAR on a large number of compounds and to increase the potency of a lead. Regardless of the application, successful use of combinatorial libraries is highly dependent on the sensitivity and specificity of the assays that are used to identify and characterize ligands.

In this section, the various combinatorial library methods will be discussed in terms of the assays that are used. The assay formats are closely matched to the mode of presentation of the diversity. In broad terms, assay procedures can be grouped into three categories: (i) those that rely on affinity purification with an immobilized target receptor, (ii) those in which a soluble receptor binds to tethered ligands, and (iii) those in which soluble compounds are tested for activity, either directly or in competition assays. Each format presents different challenges with regard to the minimum affinity requirements for ligand detection, the demonstration of binding specificity, and the ability to discriminate among compounds in the library on the basis of their affinities for the target.

Isolation of Ligands by Affinity Purification

Recombinant Peptide Libraries. The various systems described in the first part of this series[1] for creating vast libraries of recombinant peptides (commonly referred to as peptide/nucleic acid complexes below) rely on affinity purification to select peptides that bind to a receptor. Two distinct methods have been employed to achieve affinity purification of peptide/nucleic acid complexes. The first involves incubation of a receptor in solution with the mixture of complexes. After allowing sufficient time for binding, the receptor is captured using immobilized streptavidin or an antireceptor antibody.[43,44] The second approach calls for preimmobilization of the receptor on beads, microtiter wells, or a chromatography support, followed by capture of the complexes.[45] In both cases, the use of a solid-support facilitates the separation and washing of receptor-bound complexes.

The method of receptor immobilization is a critical aspect of the successful use of recombinant peptide libraries. Because of the tremendous levels of ligand enrichment attainable through multiple rounds of selection and amplification, peptides that bind to any component of the solid support can be isolated from libraries. Peptides binding to streptavidin,[46,47] antireceptor antibodies,[48] or peptides that exhibit inherent nonspecific binding are readily isolated. Often, screening strategies employing subtractive methods and/or blocking ligands, are necessary to enhance the selection of ligands with desired binding specificity.

To enhance the probability of isolating peptide ligands with biological function, it is important to ensure that the receptor is active (for example, capable of binding its natural ligand) when immobilized. Immobilization of receptor proteins on microtiter wells or beads can be accomplished by passive adherence, covalent attachment, biotinylation and immobilization on streptavidin, or capture with high-affinity nonblocking antibodies. The first three processes often result in inactive proteins. The problem of immobilizing active protein can often be overcome by introducing into the receptor an immobilization handle through genetic engineering techniques. Peptide epitopes for a monoclonal antibody or a motif that allows for site-specific biotinylation of the protein[49] can be fused to proteins for this purpose. Generic immobilization strategies of this type greatly facilitate the creation of a high-density affinity matrix suitable for isolation of ligands.

Some drug-discovery targets may not be readily available as pure soluble receptors. Peptide ligands specific for the integrin IIb/IIIa have been successfully isolated from phage libraries by screening against platelets expressing a high density of this receptor.[50] It remains to be seen whether whole cells bearing receptors or other forms of impure macromolecular target will generally be successfully utilized to identify ligands. It can be anticipated that the problem of isolating non-receptor-binding sequences will be enhanced when dealing with impure forms of receptors.

The choice of using preimmobilized receptor rather than incubation with receptor in solution followed by receptor capture on a solid support may become important in one aspect of library screening. Because the phage and LacI/DNA complex systems are capable of displaying multiple copies of a peptide, multivalent binding can occur during affinity purification if receptors are immobilized at sufficient density. Multivalent binding effectively increases the avidity of the bound complexes and allows the isolation of complexes bearing peptides of lower affinity. Although it is possible that multivalent interactions may occur during the capture step of the two-step procedure, the stoichiometry of the phage or LacI complex, receptor, and the capture reagent need to be carefully controlled. Multivalent binding conditions may be more easily engineered with a high-density preimmobilized receptor.

The multivalency of the phage and LacI systems can be exploited to isolate peptides of modest affinities (K_d values of 1–1000 μM). This feature is particularly important in screening random libraries of peptides greater than six or seven amino acids in length. Libraries that can be routinely made have many fewer members (10^8–10^{10}) than the theoretical number of possible sequences for a given peptide length (for example, there are 1×10^{13} possible 10-mers). In general, peptide ligands for a receptor target consist of families of related sequences with few high affinity ligands in the family. Therefore, when a library is created, it is likely that only lower affinity members of the family will be represented. The ability to identify these low-affinity ligands then permits one to proceed to the secondary phase of discovery, the screening of mutagenesis libraries.

Methods for creating many variants of an initial sequence have been described in part 1. Such libraries will generally contain many low-affinity ligands and perhaps some high-affinity ligands in much lower abundance. A demand is therefore placed on the ability to selectively isolate the highest affinity ligands. Several methods have been described for efficient affinity selection. All are based on the prevention or disruption of multivalent interactions. The use of a low density of immobilized receptor to isolate high-affinity peptide ligands from a phage library containing many low-affinity ligands has been demonstrated.[44] Low receptor density reduces the possibility of multivalent interactions between phage particles and the receptor matrix. The "monovalent-phage" approach[45,51,52] has been successfully employed to isolate high-affinity mutants of human growth hormone displayed on phage. In this approach, phage particles with only a single chimeric pIII protein are created, thereby eliminating multivalent binding to immobilized receptor. Using this approach, mutants with K_d values of less than 5 pM have been identified.

An alternative method of affinity selection that may have advantages over other approaches has been developed.[48,53] Phage (or LacI particles) are allowed to bind in a multivalent fashion to a high density of immobilized receptor. For phage-bearing low-affinity ligands, the peptide on an individual pIII protein may be rapidly dissociating and reassociating, but the phage particle will not dissociate unless all the peptides on pIII are simultaneously in the unbound state. Dissociation of the phage can be initiated by addition of a competing ligand, which prevents rebinding of any individual peptide in the complex. Using a model system with peptides of known affinity, it was demonstrated that phage-bearing high-affinity peptides are retained for a greater length of time than phage with lower affinity sequences.[48] The concentration (and affinity) of the competing ligand, as well as the time and temperature of elution, can be varied to select for ligands of various affinities. This method has the advantage of using a high receptor density to ensure a full sampling of ligands of all affinity classes.

Achieving affinity selection is only part of the process of successfully screening recombinant peptide libraries. After selection, it is necessary to establish the binding specificity and, if possible, the affinity of individual peptides that result from the selection. Various assays have been described, including dot blots,[54] colony lifts,[55] and ELISA's with immobilized phage or immobilized receptor.[44] These methods differ in the minimum ligand affinity that is required for detection. In general, assays in which phage or LacI are immobilized (ELISA's, dot blots, colony lifts) require higher affinity (K_d values < 1 μM) peptides for detecting specific binding, and are therefore useful when such ligands are present in the selected pool. However, for reasons cited above, detection of the specific binding of lower affinity ligands is often necessary. In such cases, assays that use a high density of immobilized receptor are required to allow for multivalent binding and to increase the sensitivity of detection. If high-density receptor matrices are used for affinity purification and assay of individual clones, peptides with K_d values as high as 100–500 μM can be isolated with phage and LacI systems.[48]

An additional assay format has been described for estimating the affinity of peptides displayed by individually selected phage clones.[44] Radiolabeled receptor is first allowed to bind to the phage-borne peptides in solution. A high concentration of competing peptide is then added to prevent further binding, and the dissociation of radiolabeled receptor is followed with time. With a monoclonal antibody model system, a good correlation was observed between dissociation rates and the affinity of the peptide determined by solution methods. It has also been suggested that colony lifts with limited receptor concentration may allow discrimination of individual phage clones on the basis of their peptide's affinity.[55] However, this method may be confounded by differential levels of expression of phage by different colonies.

In summary, the successful identification of ligands from recombinant random peptide libraries depends not only on the nature and size of the libraries but also on effective screening strategies. Selection methods and assays of individual clones vary in their ability to select and detect lower affinity peptides and in the ease with which binding specificity can be determined. High-affinity ligands are most desirable, and initial conditions for screening of random libraries can be adjusted so that only high-affinity ligands are selected. However, for reasons stated above,

engineering selection and assay conditions to allow isolation and detection of lower affinity ligands may be generally a more reliable strategy. These initial peptides can serve as starting points for creating secondary recombinant peptide libraries or as leads for refinement by synthetic chemical combinatorial methods.

Affinity Purification of Mixtures of Soluble Synthetic Compounds. Investigators have employed affinity purification methods to isolate ligands from mixtures of soluble peptides[26,56,57] and oligonucleotides (RNA or DNA).[58-60] In the case of nucleic acid libraries, one takes advantage of the ability to enzymatically amplify the molecules resulting from affinity purification, and as with the recombinant peptide systems, multiple rounds of selection and amplification are used. Theoretical considerations in optimizing conditions for the selection of high-affinity oligonucleotides have been described.[61] The authors illustrate (by way of computer simulation) the importance of nucleic acid and receptor concentrations as well as the efficiency of separating specifically bound molecules. Computer simulations show that, under ideal conditions, rare high-affinity molecules can be isolated from large libraries with relatively few rounds of selection and amplification. There have been a number of examples of successful identification of high-affinity oligonucleotides using this process.[62]

Relatively little work has focused on the affinity purification of ligands from soluble peptide libraries. The isolation of ligands for an anti-gp120 antibody from equimolar mixtures containing 19 or 32 peptides has been reported.[56,57] More recently, the same antibody was used to capture ligands from four mixtures, each of 50 peptides, comprised of unnatural amino acids fused to an encoding L-amino acid peptide strand.[26] After affinity purification, the resulting pool of peptides selected was resolved by HPLC and each peak subjected to Edman sequencing and mass spectrometry analysis. A major limitation of this approach is the sensitivity of these analytical methods. Sufficient peptide (>1–10 pmol) must be recovered in order to determine its sequence, requiring that each library member be present in relatively high amounts in the starting pool and that there be sufficient receptor available to isolate the requisite quantity of each of the high-affinity ligands. In addition, the method requires that selected peptide ligands be resolved chromatographically. While it seems unlikely that this methodology will be extended beyond libraries of modest size (less than a few thousand members), it may prove a useful technique for evaluating secondary (analog) libraries. The proposed approach of creating a library of soluble compounds with attached oligonucleotides tags may allow for the structural identification of minute quantities of compounds isolated by affinity purification.[63,64]

In theory, chromatography of compound mixtures using receptor columns should not only facilitate separation of nonbinding members of the library, but should also allow for the resolution of compounds on the basis of their receptor affinities. Work with various model systems has demonstrated that column retention time can be used as an index of affinity.[65,66] While columns of receptor target have been used in batch affinity purification methods, chromatography to resolve ligands of differing affinities has yet to be applied to screening combinatorial libraries. This method may be better suited to isolation of ligands of moderate affinity.[67] An additional limiting factor in the use of chromatography may be the large amount of receptor required to generate enough theoretical plates to effectively resolve compounds.

Binding of Receptors to Immobilized Ligands

Various methods for creating libraries of compounds attached to solid supports (pins, beads, chips, etc.) have been outlined in part 1.[1] Such libraries are screened by detecting the direct binding of a labeled receptor to an immobilized ligand. The identify of the ligand is either determined directly (by peptide sequencing or mass spectrometry), specified by its spatial location in an array, or deduced by reading an encoding tag.

There are a number of important issues related to solid-phase binding assays with immobilized ligands. First, the ability of a receptor to interact with a tethered ligand may be influenced by the site or nature of its covalent attachment to the solid support. In all of the methods published to date, peptide ligands are attached to a linker and support via the carboxy terminus of the sequence. An obvious example of the limitation imposed by this mode of immobilization would be in screening against the G-protein-linked receptors of various peptide hormones, many of which require a free C-terminal carboxamide for activity. In such a case, it is likely that many peptide analogs that would bind when free in solution would be missed in an assay where the same peptides were immobilized via their C-termini. To circumvent this problem, it is advantageous to have several alternative sites of ligand attachment to the surface. Methods for tethering peptides through their N-termini have been identified.[68] It is likely that the issue of how best to tether molecules to surfaces will become even more important when dealing with libraries of small nonpolymeric organic compounds.

The chemical nature of the linkage between the ligand and support may also affect the receptor-ligand interaction. One needs only to look at the variety of resins that are available for affinity chromatography to appreciate the importance of controlling the receptor–ligand interface. The types of linker groups that have been successfully employed in tethered library assays to date have been noted in part 1 of the series. Whether these linkers will generally provide for optimal presentation of compounds to other receptor systems remains to be seen.

Immobilized ligand assays require that the receptor be labeled in a way that allows for highly sensitive detection of receptor binding. The receptor can either be labeled directly or a secondary labeled reagent with high affinity for the receptor can be used. To date, colorimetric enzymes, radioisotopes, and fluorophores have been used in labeling receptors or secondary reagents. The reagents must be labeled in a way that maintains the activity of the receptor, for instance, its ability to bind a natural ligand. This can be greatly facilitated by creating chimeric recombinant receptors that incorporate peptide epitopes of antibodies or peptide sequences for site-specific radioactive phosphorylation[69] or site-specific biotinylation.[49]

Successful screening of libraries of immobilized synthetic ligands is dependent on the same types of issues as have been previously discussed with respect to evaluating recombinant peptide libraries: i.e., the affinity threshold for detection, the ability to discriminate ligands on the basis of their affinities, and the ability to distinguish specific binding from nonspecific binding. Methods development in this area is in its infancy. In principle, it

should be possible to exploit multivalent binding to detect lower affinity ligands. Multivalent receptors can be created by a number of methods, including genetic fusions to generate bivalent receptor/Fc fusions[70] or through the use of monoclonal antibodies or streptavidin to create cross-linked receptors capable of interaction with more than one immobilized ligand. Optimization of the density of immobilized ligands may be required in order to allow for multivalent binding. As has already been noted, it may be important to be able to isolate relatively low-affinity ligands in the initial screening of random libraries. These compounds can then serve as the basis for further library construction in which the goal is to improve ligand affinity.

Affinity discrimination during the screening of either primary random libraries or secondary (analog) libraries is of obvious importance. There has been little published work on methodology in this area. In principle, low receptor concentrations, competing ligand-mediated dissociation, and/or stringent washing conditions can be utilized to identify the highest affinity ligands. Two issues complicate the use of such methods. The first is the likelihood that each pin, bead, or surface synthesis site does not contain the same amount of compound. With different compound loadings, one must be extremely cautious of using the absolute quantity of bound receptor as an index of a molecule's affinity. As new building-block and coupling chemistries are adapted to combinatorial formats, this may become a more significant problem than it is for high-yielding peptide chemistry. Another complicating feature of the immobilized ligand assay format is the fact that ligands of one particular kind are densely clustered on a surface. Both the association and dissociation rate constants of a receptor/ligand interaction are affected by surface ligand density. The binding of nearby ligands depletes the local receptor concentration and the association kinetics become diffusion limited. After dissociation, receptor rebinding is favored because of the high local-ligand concentration and the apparent dissociation rate is reduced. Theoretical and experimental analyses of these surface binding effects have been undertaken.[71] The impact that these surface binding kinetics will have on the ability to discriminate among library members on pins, beads, or glass surfaces remains to be seen.

The information generated by screening immobilized ligand libraries differs among the various library formats. In the case of bead-based technologies, compounds exceeding a threshold affinity are sampled from a large pool of ligands. Positive information is obtained, i.e., that a particular ligand binds to the receptor. One cannot, however, draw conclusions about the binding affinity of nonselected ligands. The sampling size may not have been large enough to include all high-affinity ligands, or a high-affinity bead may have been missed by the affinity selection method [for example, fluorescence-activated cell sorter (FACS) selection]. By contrast, the multipin and VLSIPS technologies allow one to perform a parallel assay in which data is obtained on every compound that is synthesized. In principle, both positive and negative binding information can be exploited in the design of second-generation compounds.

Incorporation of methods that assess the specificity of binding of ligands is an important aspect of screening random libraries. Screening immobilized ligands by direct receptor binding can lead not only to the identification of ligands of interest (for instance, ligands that compete with the natural ligand) but also to ligands that bind to undesired portions of the receptor or to secondary detection reagents. In the case of libraries of compounds on beads, it may be possible to remove undesired ligands in a subtraction step prior to screening for desired ligands. For compounds on pins or chips, it may be possible to make replicate arrays and test for total binding and nonspecific binding in parallel. Otherwise, sequential assays that first test for receptor binding of any kind, followed by an assessment of nonspecific binding will be required in order to correctly identify compounds that interact with the receptor in a desired manner.

Testing the Activity of Libraries of Soluble Compounds

The classical method of screening for a desired biological activity is to test soluble compounds one at a time in a competition binding assay, enzyme inhibition assay, or in a cell-based bioassay. Such approaches have been applied to library screening by releasing compounds synthesized on pins into microtiter wells, as described in section C of part 1 of this series.[1] A novel application of bead technology has recently been disclosed where compounds on individual beads are released locally onto a lawn of confluent mammalian cells and cause activation of cells in the area surrounding the bead.[72] The bead responsible for cell activation is isolated and a small amount of noncleaved peptide is sequenced to determine its structure. In both of these cases, the principal issue is whether enough compound is released to be detected by the assay. For pins, approximately 100 nmol of peptide can be released into a few hundred microliters of solution, while beads with diameters of $\sim$100 μm can release on the order of 100 pmol of peptide.

Rather than assaying compounds individually, a second approach to screening soluble libraries is to assay compound mixtures. In addition to testing complex pools of soluble peptides (*vide supra*), libraries of oligonucleotides have been successfully screened as soluble mixtures.[73] The most frequently used strategy for screening mixtures of soluble compounds with the goal of ultimately identifying single active molecules is based on the "mimotope" approach, detailed in part 1. The essence of this strategy is that degenerate pools of peptides (or other compounds) are resolved into their most active constituents by an iterative process of testing and resynthesis until a single sequence is identified as having high activity. A variation of the methodology (termed "bogus-coin strategy") has also been described.[74]

There are a number of caveats to using this methodology for testing soluble compound mixtures. In practice, the results of each set of assays do not typically indicate a preference for a unique residue at any position within the sequence. Rather, comparable assay results may be obtained for several different amino acid substitutions and some decision must be made as to which of these partial solutions should be fully resolved. The number of peptide mixtures to be synthesized and tested in this protocol expands dramatically as the number of alternative sequences selected for complete resolution at each cycle is increased. Moreover, the deconvolution of different partial solutions may frequently produce divergent resolved sequences, in part because the contribution of each amino acid to the peptide–receptor interaction is typically

dependent on other non-neighboring residues within the ligand. The problem of identifying the most potent ligand in a complex mixture by this type of iterative pathway is exacerbated by the relative abundance of lower affinity ligands that represent local binding optima.

Originally designed for identifying antibody ligands, the mimotope strategy has primarily been used for libraries of six to eight building blocks in length. It is not clear that ligands of this size will be optimal for other types of receptors (although success with opioid receptors[75] and other targets have been reported). As the length of the compounds in the library increases, resynthesis and testing of pools becomes more cumbersome.

Perhaps the greatest limitation of this methodology is the fact that the activity of a given pool is based on the cumulative activity of all the compounds in the pool; i.e., pools with the same activity may contain many low-affinity compounds or a few high-affinity compounds. For this reason, the methodology is greatly facilitated if the minimal fragment having activity is comprised of the same number of building blocks as used in constructing each library member (e.g., a uniquely active tetramer is more easily resolved from a tetrapeptide library than a hexamer library). Alternatively, the identification of active peptide(s) is facilitated if the receptor has specific requirements for a fixed position within a peptide ligand (e.g., the N or C termini). If neither of these conditions is true, it may be necessary to test many or all of the possible initial pools with two adjacent or nonadjacent fixed residues. This drastically increases the number of initial pools that need to be synthesized but increases the probability that a critical residue(s) is fixed in at least one pool to allow that pool to differentiate itself. It must be kept in mind that any pool identified as having the greatest activity may be composed of many moderately active compounds and that the most active compound(s) may reside in other pools.

Testing of mixtures of soluble compounds is also limited by the concentration of individual test compounds that can be achieved in the initial pools. Pools containing as many as 160 000 different peptides have been tested with each member being present at ~10 nanomolar concentration.[76] Because of limitations on the solubility of the total pool, the concentrations of individual compounds present in increasingly larger libraries must be correspondingly diminished. This will ultimately limit the ability to identify the activity of compounds with modest potencies.

While the current methods for testing mixtures of soluble compounds have certain drawbacks, screening soluble libraries does have the decided advantage of avoiding the problems associated with assaying tethered molecules in other combinatorial technologies. Conventional binding and enzyme and cell-based assays (including those with poorly defined biochemical targets) can be used to test the activity of soluble compounds. It is likely that in the future, encoding strategies will be employed to allow more facile screening of soluble molecules. In the simplest format, single encoded beads can be dispensed into microtiter wells. The compounds can then be released from the beads and tested for activity, with the identity of the most active compound(s) being deduced by decoding the tag attached to the bead(s). To test large libraries of soluble compounds, mixed pools of encoded beads can be created. At each round of testing, only a fraction of the compound is cleaved from each bead. Active pools of beads are pursued by further subdividing the beads, partially releasing the compounds, and testing for activity. In the end, a single bead is identified with the greatest activity and the encoding tag is then read.[77]

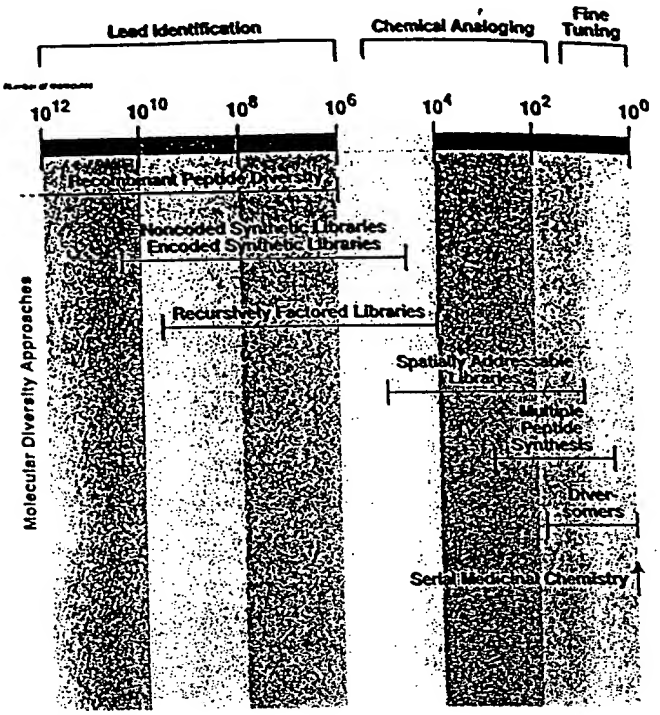


Figure 19.

C. Integration of Combinatorial Technologies for Drug Discovery and Future Directions of the Field

Regardless of whether the objective is a broad discovery search or analoging a known lead, a key aspect in the successful application of combinatorial technologies to drug discovery is the requirement for having a closely linked, coordinated process for the integration of synthesis and screening. The creation and evaluation of molecular diversity are two sides of the same coin. It is still premature to speculate on which type of format will prove most suitable for a particular type of discovery/development problem. Over time, a cataloging of successes and failures will clarify this issue. In all likelihood, command of a collection of combinatorial tools will be required for general success. One may imagine a "spectrum of molecular diversity" stretching from few to many molecules (see Figure 19). Serial synthesis (contemporary medicinal chemistry) operates on a few molecules (far right of chart). We expect that each individual combinatorial tool/format will turn out to be most valuable within proscribed regions of numbers of molecules. Hence, recombinant peptide diversity is particularly suitable for generating and screening large libraries of $>10^8$ compounds. The VLSIPS chip technology, although capable of displaying vast arrays, is primarily an analoging tool and is most useful for evaluating 10^2–10^4 compounds. The parallel organic synthesis methods of Hobbs DeWitt *et al.*[36] are applicable for tens of compounds. Encoded synthetic libraries appear to be relevant in broad diversity searching and may also prove useful in narrower optimization strategies. Given the repetitive nature of many of the manipulations required for library construction, an on-going priority will be to address the possibility of automating as many aspects of the generation/evaluation process as possible. Growth in

library size also impacts directly on the physical size of compound collections and on the amounts of target receptor required for screening. These pressures will continue to drive the field toward minaturization and exploitation of nanotechnologies.

The power of combinatorial technologies in generating huge numbers of compounds suggests that in a lead-discovery mode, less preconceived bias need be brought to the process of making molecules. Another way of expressing this is as follows: let the numbers do the talking. Due to the time and effort required in serial approaches, each target molecule is selected with great care. Because of the relative ease in creating libraries, little risk is incurred, or effort expended, in allowing a wide variety of building blocks to participate in diversity generation. Since there is less up-front investment in any individual combinatorially created molecule, the combinatorial chemist can afford to take more risks. We can think in terms of a portfolio of libraries which might be routinely applied to the initiation of a drug-discovery search. This is not dissimilar to selecting a preliminary screening sample of diversity from a large database of individual molecules.

A related, but still immature, issue in combinatorial approaches to drug discovery revolves around the idea of "quantitation of diversity". An understanding of the concept of "measuring" molecular diversity could impact on designing libraries to contain maximal structural diversity. This notion has arisen previously in deciding which few representative, highly diverse compounds to select out of large database collections, when setting up groups of preliminary screening samples. The huge numbers involved in combinatorial approaches intensifies this issue. A number of interesting approaches to the diversity quantitation problem can be expected to emerge.

One working drug-discovery paradigm might be based on initially employing a portfolio of biological diversity (peptide libraries) together with standard chemical libraries (various-sized cyclic peptides, cysteinyl-linked cyclics, etc.), peptides with carboxyl or carboxamide display, synthetic polymeric diversity, as well as large libraries of semirigid and acyclic small molecules prepared by COS. Over time, favored libraries and directions would be expected to emerge. As the sophistication of combinatorial organic synthesis grows, the origins of a molecular structure as either "combinatorially or serially derived" will gradually become transparent.

Another area where considerable effort must be applied is in the registry of libraries and individual library members. It is unclear that library compounds should be registered and documented for testing in the same ways as serially produced compounds historically have been, but exactly what changes are necessary remain to be determined. Vast numbers of compounds have been and are being created; keeping track of these and their corresponding biological activities will require innovative database-management techniques. Additionally, nomenclature needs to be developed by which one can simply express the constitution, scope, and nature of chemical libraries. Legal issues, including the patenting and documentating of libraries and their component members, will need to be pioneered.

As repeatedly emphasized, it is obvious but imperative to have efficient means of evaluating the molecular diversity which is generated. Different assay techniques will be format specific. Assays must clearly discriminate specific from nonspecific binding. Since in a broad screening mode, one is almost always sampling a small percentage of the entire universe of diversity (10^{10} peptides are only 0.1% of the universe of 10-mers), it is crucial that appropriate assay techniques be competent to detect modest affinity ligands. The identification of weak binders in any of the aforementioned approaches is very important and should lead directly to preparation of secondary libraries in which original "hits" will become the centerpiece for more focused diversity creation. This is a consequential issue, since application of combinatorial technologies are best viewed as an iterative process and not a singular event. As the emphasis shifts to analog evaluation, assays must be capable of affinity discrimination between closely related library members. The tools of molecular biology have permitted the molecular engineering of targets to serve the purposes of screening. The rapid introduction of targets into a screening mode will require generic techniques for their handling, and manipulation of molecular targets by genetic engineering will continue to play a crucial role in marrying library evaluation and synthesis. Though combinatorial technologies may soon prove their worth in the drug-discovery process by delivering new leads quickly and cheaply, in order to completely fulfill the promise of "making drugs", an important question will be whether some of the common major obstacles to drug development (e.g., cell penetration, bioavailability, pharmacokinetics, metabolism) can be productively addressed through the application of combinatorial approaches (i.e., *combinatorial drug development*).

In the coming years, cloning and sequencing of the human genome promises that an unprecedented abundance of newly discovered proteins will become available as potential drug targets. Gaining even more prominence than it now assumes will be the issue of discriminating among a myriad of receptors and enzymes to identify valid targets for drug discovery. The ability to access potent and specific ligands for these targets will guide this process by untangling the physiological relevance of endogenous biochemical pathways. Combinatorial methods will be called upon to provide such molecules to quickly and cheaply drive target validation. In this manner, the identification of leads will benefit from a significant, but hidden, benefit which emerges from combinatorial screening; hits derived from chemical libraries should be readily amenable to combinatorial analoging.

Certain drug targets may present more or less of a historical precedent with respect to the likelihood of successfully identifying a tight binding ligand through the use of known pharmacophores. For example, the search for specific enzyme inhibitors may be facilitated by the intentional enrichment in the combinatorial synthetic process of building blocks containing known inhibitory pharmacophores. Particularly important or common types of drug targets may justify having on hand special libraries which are somewhat specific (i.e., a peptidyl hydroxyethylamine library for aspartyl- and metalloprotease inhibition[78-80]). On the other hand, in areas where there is less current information (e.g., antagonism of protein–protein or carbohydrate–protein interactions), a wider scope of diversity search should be taken until consistent patterns begin to emerge. In the case of newer, less explored target groups, combinatorial technologies can be expected to assist in unearthing new pharmacophore

classes and to help establish an understanding of drug design for new types of targets.

Combinatorial technologies diverge sharply from historical precedent through a change in emphasis from the consideration of individual molecules to thinking in terms of populations of molecules. A common, but false, intuitive believe is that combinatorial chemistry is necessarily a random, screening search; the antithesis of rational drug design. In fact, all libraries are biased in some ways. All drug company compound files are biased by the historical programs of that institution, since a disproportionate share of compounds of particular types will have been deposited. *The notion of intentionally biasing a chemical library is a form of drug design*, but again not applied to individuals but rather to groups or populations of molecules. If a scientist hypothesizes on the basis of structural information that the current lead molecule contains a type II β-turn motif, then rather than performing two or three serial tests of this idea, the combinatorial chemist might create a library of narrow diversity utilizing a basis set of β-turn mimetics and thus interrogate many slightly different regions of conformational space simultaneously. The drug design of populations versus individuals is analogous to fishing with a net rather than just a hook. As more knowledge of workable strategies for combinatorial synthesis are understood, it is expected that structural and computational input and other rational design information will be integrated into a broad combinatorial medicinal chemistry approach.

Gaining a full appreciation of the issues and difficulties which must be surmounted in order to perform useful combinatorial organic synthesis will initially be a relatively slow process, especially because the important strategies and decision points differ so markedly from traditional organic synthesis. Retrocombinatorial analysis of existing pharmacophores and other important structures should assist in decision making; both in choosing routes of forward synthesis and in synthetic target selection. If combinatorial techniques are indeed to become a useful shortcut to new leads and optimized compounds, then one key implied goal of combinatorial organic synthesis is to intersect the pathway of modern medicinal chemistry upon which compounds move from the early discovery stages to clinical candidacy. Rich incentives await those who are able to mass produce important biologically active molecules quickly and cheaply. Not surprisingly, an aggressive, worldwide effort to understand and master this field has already begun.

This Perspective has been restricted to a consideration of the impact of combinatorial technologies on medicinal chemistry/drug discovery and development. From the point of view of applicability of the technologies, this is an artificially narrow view. Combinatorial processes will become important in diagnostic medicine,[81] agricultural chemistry, food chemistry, immunology, molecular biology, polymer studies, inorganic synthesis, and many other fields. Though the field of "combinatorial chemistry" is chronologically a new enterprise, the evolution of thought in this fertile area continues to outrace the experimental reduction to practice of many ideas. One may reasonably ask "why are combinatorial technologies happening now?". The answer is probably complex and beyond the scope of this Perspective. Nevertheless, the explosive recent interest in the application of combinatorial technologies to drug discovery is symptomatic of an idea whose time has come. Because the issues which confront the medicinal chemist differ so radically from historical approaches, the combinatorial field will no doubt continue to provide impetus and stimulation for the formulation of new concepts and ideas.

Acknowledgment. We thank Drs. A. Zaffaroni and G. Ringold for their interest, insights, and support in undertaking the application of combinatorial technologies to drug discovery and Mrs. Terri Nebozuk for help in preparing this manuscript.

References

(1) Part 1: Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P. A.; Gordon, E. M. Applications of Combinatorial Technologies to Drug Discovery. 1. Background and Peptide Combinatorial Libraries. *J. Med. Chem.* 1994, *37*, 1233–1251.
(2) Barker, P. L.; Bullens, S.; Bunting, S.; Burdick, D. J.; Chan, K. S.; Deisher, T.; Eigenbrot, C.; Gadek, T. R.; Lipari, M. T.; *et al.* Cyclic RGD Peptide Analogues as Antiplatlet Antithrombotics. *J. Med. Chem.* 1992, *35*, 2040–2048.
(3) Al-Obeidi, F.; De L Castrucci, A. M.; Hadley, M. E.; Hruby, V. J. Potent and Prolonged Acting Cyclic Lactam Analogues of an α-Melanotropin: Design Based on Molecular Dynamics. *J. Med. Chem.* 1989, *32*, 2555–2561.
(4) Wiley, R. A.; Rich, D. H. Peptidomimetics Derived from Natural Products. *Med. Res. Rev.* 1993, *13*, 327–384.
(5) Giannis, A.; Kolter, T. Peptidomimetics for Receptor Ligands—Discovery, Development, and Medicinal Perspectives. *Angew. Chem., Int. Ed. Engl.* 1993, *32*, 1244–1267.
(6) Kates, S. A.; Sole, N. A.; Johnson, C. R.; Hudson, D.; Barany, G.; Albericio, F. A Novel, Convenient, Three-Dimensional Orthogonal Strategy for Solid Phase Synthesis of Cyclic Peptides. *Tetrahedron Lett.* 1993, *34*, 1549–1552.
(7) Polinsky, A.; Cooney, M. G.; Toy-Palmer, A.; Osapay, G.; Goodman, M. Synthesis and Conformational Properties of the Lanthionine Bridged Opioid Peptide [D-Ala$_2$,Ala$_5$]Enkephalin as determined by NMR and Computer Simulations. *J. Med. Chem.* 1992, *35*, 4185–4194.
(8) Albericio, F.; Hammer, R. P.; Garcia-Echeverria, C.; Molins, M. A.; Chang, J. L.; Munson, M. C.; Pons, M.; Giralt, E.; Barany, G. Cyclization of Disulfide-Containing Peptides in Solid Phase Synthesis. *Int. J. Pept. Protein Res.* 1991, *37*, 402–413.
(9) Marlowe, C. K. Peptide Cyclization on TFA Lable Resin Using the Trimethylsilyl (TMSE) Ester as an Orthogonal Protecting Group. *Bioorg. Med. Chem. Lett.* 1993, *3*, 437–440.
(10) Trzeciak, A.; Bannwarth, W. Synthesis of 'Head to Tail' Cyclized Peptides on Solid Support by FMOC Chemistry. *Tetrahedron Lett.* 1992, *33*, 4557–4560.
(11) Andrews, P. C.; Boyd, J.; Ogorzalek Loo, R.; Zhao, R.; Zhu, C.-Q.; Grant, K.; Williams, S. Synthesis of Uniform Peptide Libraries and Methods for Physico-Chemical Analysis. Poster presented at the Meeting of the Protein Society, San Diego, July 1993.
(12) Metzger, J. W.; Wiesmuller, K.-H.; Gnau, V.; Brunjes, J.; Jung, G.; Ion-Spray Mass Spectrometry and High Performance Liquid Chromatography–Mass Spectrometry of Synthetic Peptide Libraries. *Angew. Chem., Int. Ed. Engl.* 1993, *32*, 894–896.
(13) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C. Design, Construction and Application of a Fully Automated Equimolar Peptide Mixture Synthesizer. *Int. J. Pept. Protein Res.* 1992, *40*, 497–506.
(14) Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. Light-Directed, Spatially Addressable Parallel Chemical Synthesis. *Science* 1991, *251*, 767–773.
(15) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C.; Santi, D. V. Identification of Highest-Affinity Ligands by Affinity Selection from Equimolar Peptide Mixtures Generated by Robotic Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 4505–4509.
(16) Freidinger, R. M.; Perlow, D. S.; Veber, D. F. Protected Lactam-Bridged Dipeptides for Use as Conformational Constraints in Peptides. *J. Org. Chem.* 1982, *47*, 104–109.
(17) Crowley, J. I.; Rapoport, H. Solid-Phase Organic Synthesis: Novelty or Fundamental Concept. *Acc. Chem. Res.* 1976, *9*, 135–144.
(18) Leznoff, C. C. The Use of Insoluble Polymer Supports in General Organic Synthesis. *Acc. Chem. Res.* 1978, *11*, 327.
(19) Frechet, J. M. J. Synthesis and Applications of Organic Polymers as Supports and Protecting Groups. *Tetrahedron* 1981, *37*, 663–683.
(20) Hodge, P. Organic Reactions Using Polymer-Supported Catalysts, Reagents or Substrates. In *Syntheses and Separations Using Functional Polymers*; Sherrington, D. C., Hodge, P., Eds.; John Wiley and Sons: New York, 1988, pp 43–122.
(21) Cho, C. Y.; Moran, E. J.; Cherry, S. R.; Stephans, J. C.; Fodor, S. P. A.; Adams, C. L.; Sundaram, A.; Jacobs, J. W.; Schultz, P. G. An Unnatural Biopolymer. *Science* 1993, *261*, 1303–5.

(22) Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K.; Spellmeyer, D. C.; Tan, R.; Frankel, A. D.; Santi, D. V.; Cohen, F. E.; Bartlett, P. A. Peptoids: A Modular Approach to Drug Discovery. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 9367–9371.
(23) Bartlett, P. A.; Santi, D. V.; Simon, R. J. Int. Patent WO91/19735, 1991.
(24) Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. Efficient method for the preparation of peptoids[oligo(N-substituted glycines)] by submonomer solid-phase synthesis. *J. Am. Chem. Soc.* 1992, *114*, 10646–10647.
(25) Zuckermann, R. N. Presented at the Third International Symposium on Solid Phase Synthesis and Complementary Technologies, Oxford, U.K., 31 August–4 September, 1993.
(26) Kerr, J. M.; Banville, S. C.; Zuckermann, R. N. Encoded Combinatorial Peptide Libraries Containing Non-Natural Amino Acids. *J. Am. Chem. Soc.* 1993, *115*, 2529–2531.
(27) Rich, D. H. Peptidase Inhibitors. In *Comprehensive Medicinal Chemistry*; Hansch, C., Sammes, P. G., Taylor, J. B., Eds.; Pergamon Press: Oxford, 1990, Vol. 2.
(28) Bartlett, P. A.; Marlowe, C. K. Possible Role for Water Dissociation in the Slow Binding of Phosphorous-Containing Transition-State-Analogue Inhibitors of Thermolysin. *Biochemistry* 1987, *26*, 8553–8561.
(29) Morgan, B. P.; Scholtz, J. M.; Ballinger, M. D.; Zipkin, I. D.; Bartlett, P. A. Differential Binding Energy: A Detailed Evaluation of the Influence of Hydrogen-Bonding and Hydrophobic Groups on the Inhibition of Thermolysin by Phosphorous-Containing Inhibitors. *J. Am. Chem. Soc.* 1991, *113*, 297–307.
(30) Campbell, D. A.; Bermak, J. C. Phosphonate Ester Synthesis Using a Modified Mitsunobu Condensation. *J. Org. Chem.* 1994, *59*, 658–660. Campbell, D. A. The Synthesis of Phosphonate Esters, an Extension of the Mitsunobu Reaction. *J. Org. Chem.* 1992, *57*, 6331–6335.
(31) Hagihara, M.; Anthony, N. J.; Stout, T. J.; Clardy, J.; Schreiber, S. L. Vinylogous Polypeptides: An Alternative Peptide Backbone. *J. Am. Chem. Soc.* 1992, *114*, 6568–6570.
(32) Smith, A. B., III; Keenan, T. P.; Holcomb, R. C.; Sprengler, P. A.; Guzman, M. C.; Wood, J. L.; Carroll, P. J.; Hirschmann, R. Design, Synthesis and Crystal Structure of a Pyrrolinone-Based Peptidomimetic Possessing the Conformation of a β-Strand: Potential Applications to the Design of Novel Inhibitors of Proteolytic Enzymes. *J. Am. Chem. Soc.* 1992, *114*, 10672–10674.
(33) Bunin, B. A.; Ellman, J. A. A General and Expedient Method for the Solid-Phase Synthesis of 1,4-Benzodiazepine Derivatives. *J. Am. Chem. Soc.* 1992, *114*, 10997–10998.
(34) Ellman, J. A. Pharmaceutical Manufacturers Association, Drug Discovery Management Subsection, September 19–21, 1993, Philadelphia, PA.
(35) Nikolaiev, V.; Stierandova, A.; Krchnak, V.; Seligmann, B.; Lam, K. S.; Salmon, S. E.; Lebl, M. Peptide-Encoding for Structure Determination of Nonsequenceable Polymers Within Libraries Synthesized and Tested on Solid-Phase Supports. *Pept. Res.* 1993, *6*, 161–170.
(36) Hobbs DeWitt, S.; Kiely, J. S.; Stankovic, C. J.; Schroeder, M. C.; Reynolds Cody, D. M.; Pavia, M. R. "Diversomers": An Approach to Nonpeptide, Nonoligomeric Chemical Diversity. *Proc. Natl. Acad. Sci. U.S.A.* 1993, *90*, 6909–6913.
(37) Mapelli, C.; Swerdloff, M. D. Monitoring of Conformational and Reaction Events in Resin-Bound Peptides by Carbon-13 NMR Spectroscopy in Various Solvents. In *Peptides. Proceedings of the 21st European Peptide Symposium*; Giralt, E. and Andreu, D., Eds., ESCOM, Leiden, The Netherlands, 1990; pp 316–319.
(38) Giralt, E.; Rizo, J.; Pedroso, E. Application of Gel-Phase ^{13}C-NMR to Monitor Solid Phase Peptide Synthesis. *Tetrahedron* 1984, *40*, 4141–4152.
(39) Nicolaou, K. C.; Salvino, J. M.; Raynor, K.; Pietranico, S.; Reisine, T.; Freidinger, R. M.; Hirschmann, R. Design and Synthesis of a Peptidomimetic Employing β-D-Glucose for Scaffolding. *Peptides* 1989, 881–884.
(40) Hirschmann, R.; Nicolaou, K. C.; Pietranico, S.; Salvino, J.; Leahy, E. M.; Sprengeler, P. A.; Furst, G.; Smith, A. B.; III; Strader, C. D.; Cascieri, M. A.; Candelore, M. R.; Donaldson, C.; Vale, W.; Maechler, L. Nonpeptidal Peptidomimetics with a β-D-Glucose Scaffolding. A Partial Somatostatin Agonist Bearing a Close Structural Relationship to a Potent, Selective Substance P Antagonist. *J. Am. Chem. Soc.* 1992, *114*, 9217–9218.
(41) Hirschmann, R.; Sprengeler, P. A.; Kawasaki, T.; Leahy, J. W.; Shakespeare, W. C.; Smith, A. B., III, The First Design and Synthesis of a Steroidal Peptidomimetic. The Potential Value of Peptidomimetics in Elucidating the Bioactive Conformation of Peptide Ligands. *J. Am. Chem. Soc.* 1992, *114*, 9699–9701.
(42) Hirschmann, R.; Sprengeler, P. A.; Kawasaki, T.; Leahy, J. W.; Shakespeare, W. C.; Smith, A. B., III, The Versatile Steroid Nucleus: Design and Synthesis of a Peptidomimetic Employing this Novel Scaffold. *Tetrahedron* 1993, *49*, 3665–3676.
(43) Parmley, S. F.; Smith, G. P. Antibody-Selectable Filamentous Fd Phage Vectors: Affinity Purification of Target Genes. *Gene* 1988, *73*, 305–318.
(44) Barrett, R. W.; Cwirla, S. E.; Ackerman, M. S.; Olson, A. M.; Peters, E. A.; Dower, W. J. Selective Enrichment and Characterization of High Affinity Ligands from Collections of Random Peptides on Filamentous Phage. *Anal. Biochem.* 1992, *204*, 357–364.
(45) Bass, S.; Green, R.; Wells, J. A. Hormone Phage: An Enrichment Method for Variant Proteins with Altered Binding Properties. *Proteins: Struct. Funct. Genet.* 1990, *8*, 309–314.
(46) Devlin, J. J.; Panganiban, L. C.; Devlin, P. E. Random Peptide Libraries: A Source of Specific Protein Binding Molecules. *Science* 1990, *249*, 404–406.
(47) Kay, B. K.; Adey, N. B.; He, Y.-S.; Manfredi, J. P.; Mataragnon, A. H.; Fowlkes, D. M. An M13 Phage Library Displaying Random 38-Amino-Acid Peptides as a Source of Novel Sequences with Affinity to Selected Targets. *Gene* 1993, *128*, 59–65.
(48) Barrett, R. W.; Dower, W. J. Unpublished.
(49) Schatz, P. J. Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: A 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. *Bio/Tech.* 1993, *11*, 1138–1143.
(50) Doyle, M. Chiron Corporation. Presented at San Francisco Bay Area Peptide Society Meeting, December 1992.
(51) Lowman, H. B.; Bass, S. H.; Simpson, N.; Wells, J. A. Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry* 1991, *30*, 10832–10838.
(52) Garrard, L. J.; Yang, M.; O'Connell, M. P.; Kelley, R. F.; Henner, D. J. Fab Assembly and Enrichment in a Monovalent Phage Display System. *Bio/Technology* 1991, *9*, 1373–1377.
(53) Hawkins, R. E.; Russell, S. J.; Winter, G. Selection of Phage Antibodies by Binding Affinity: Mimicking Affinity Maturation. *J. Mol. Biol.* 1992, *226*, 889–896.
(54) Kang, A. S.; Barbas, C. F.; Janda, K. K.; Benkovic, S. J.; Lerner, R. A. Linkage of Recognition and Replication Functions by Assembling Combinatorial Antibody Fab Libraries Along Phage Surfaces. *Proc. Natl. Acad. Sci. U.S.A.* 1991, *88*, 4363–4366.
(55) Christian, R. B.; Zuckermann, R. N.; Kerr, J. M.; Wang, L.; Malcolm, B. A. Simplified Methods for Construction, Assessment and Rapid Screening of Peptide Libraries in Bacteriophage. *J. Mol. Biol.* 1992, *227*, 711–718.
(56) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C.; Santi, D. V. Identification of Highest-Affinity Ligands by Affinity Selection from Equimolar Peptide Mixtures Generated by Robotic Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 4505–4509.
(57) Kerr, J. M.; Banville, S. C.; Zuckermann, R. N. Identification of Antibody Mimotopes Containing Non-Natural Amino Acids by Recombinant and Synthetic Peptide Library Affinity Selection Methods. *Bioorg. Med. Chem. Lett.* 1993, *3*, 463–468.
(58) Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* 1990, *249*, 505–510.
(59) Ellington, A. D.; Szostak, J. W. *In vitro* Selection of RNA molecules That Bind Specific Ligands. *Nature* 1990, *346*, 818–822.
(60) Sherman, M. I.; Bertelsen, A. H.; Cook, A. F. Protein Epitope Targeting: Oligonucleotide Diversity and Drug Discovery. *Bioorg. Med. Chem. Lett.* 1993, *3*, 469–475.
(61) Irvine, D.; Tuerk, C.; Gold, L. Selexion: Systematic Evolution of Ligands by Exponential Enrichment with Integrated Optimization by Non-Linear Analysis. *J. Mol. Biol.* 1991, *222*, 739–761.
(62) Tuerk, C.; MacDougal, S.; Gold, L. RNA Pseudoknots that Inhibit Human Immunodeficiency Virus Type 1 Reverse Transcriptase. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 6988–6992.
(63) Brenner, S.; Lerner, R. A. Encoded Combinatorial Chemistry. *Proc. Natl. Acad. Sci. U.S.A.* 1992, *89*, 5181–5183.
(64) Nielsen, J.; Brenner, S.; Janda, K. D. Synthetic Methods for the Implementation of Encoded Combinatorial Chemistry. *J. Am. Chem. Soc.* 1993, *115*, 9812–9813.
(65) Arnold, F. H.; Schofield, S. A.; Blanch, H. W. Analytical Affinity Chromatography: I. Local Equilibrium Theory and the Measurement of Association and Inhibition Constants. *J. Chromatogr.* 1986, *355*, 1–12.
(66) Arnold, F. H.; Blanch, H. W. Analytical Affinity Chromatography: II. Rate Theory and the Measurement of Biological Binding Kinetics. *J. Chromatogr.* 1986, *355*, 13–27.
(67) Zopf, D.; Ohlson, S. Weak-Affinity Chromatography. *Nature* 1990, *346*, 87–88.
(68) Holmes, C. P. Polymer Reversal on Solid Surfaces. U.S. Pat. 5,242,974.
(69) Li, B.-L.; Langer, J. A.; Schwartz, B.; Petska, S. Creation of Phosphorylation Sites in Proteins: Construction of a Phosphorylatable Human Interferon α. *Proc. Natl. Acad. Sci. U.S.A.* 1989, *86*, 558–562.
(70) Hollenbaugh, D.; Chalupny, N. J.; Aruffo, A. Recombinant Globulins: Novel Research Tools and Possible Pharmaceuticals. *Curr. Opinion Immunol.* 1992, *4*, 216–219.
(71) Erickson, J.; Goldstein, B.; Holowka, D.; Baird, B. The Effect of Receptor Density on the Forward Rate Constant for Binding of Ligands to Cell Surface Receptors. *Biophys. J.* 1987, *52*, 657–662.
(72) Jayawickreme, C. K.; Graminski, G. F.; Quillan, J. M.; Lerner, M. R. Creation and Functional Screening of a Multi-Use Peptide Library. *Proc. Natl. Acad. Sci. U.S.A.* 1994, *91*, 1614–1618.

(73) Ecker, D. J.; Vickers, T. A.; Hanecak, R.; Driver, V.; Anderson, K. Rational Screening of Oligonucleotide Combinatorial Libraries for Drug Discovery. *Nucleic Acids Res.* **1993**, *21*, 1853–1856.
(74) Blake, J.; Litzi-Davis, L. Evaluation of Peptide Libraries: An Iterative Strategy to Analyze the Reactivity of Peptide Mixture with Antibodies. *Bioconj. Chem.* **1992**, *3*, 510–513.
(75) Dooley, C. T.; Chung, N. N.; Schiller, P. W.; Houghten, R. A. Acetalins: Opioid Receptor Antagonists Determined Through the Use of Synthetic Peptide Combinatorial Libraries. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10811–10815.
(76) Houghten, R. A.; Appel, J. R.; Blondelle, S. E.; Cuervo, J. H.; Dooley, C. T.; Pinilla, C. The Use of Synthetic Peptide Combinatorial Libraries for the Identification of Bioactive Peptides. *BioTechniques* **1992**, *13*, 412–421.
(77) Salmon, S. E.; Lam, K. S.; Lebl, M.; Kandola, A.; Khattri, P. S.; Wade, S.; Patek, M.; Kocis, P.; Krchnak, V.; Thorpe, D.; Felder, S. Discovery of Biologically Active Peptides in Random Libraries: Solution-Phase Testing After Staged Orthogonal Release from Resin Beads. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11708–11712.
(78) Gordon, E. M.; Godfrey, J. D.; Pluscec, J.; Von Langen, D.; Natarajan, S. Design of Peptide Derived Amino Alcohols as Transition-State Analog Inhibitors of Angiotensin Converting Enzyme. *Biochem. Biophys. Res. Commun.* **1985**, *126*, 419–426.
(79) Godfrey, J. D.; Gordon, E. M.; Von Langen, D.; Engebrecht, J. Synthesis of Peptidyl Derived Amino Alcohols. *J. Org. Chem.* **1986**, *51*, 3073.
(80) Rich, D. H.; Green, J.; Toth, M. V.; Marshall, G. R.; Kent, S. B. H. Hydroxyethylamine Analogs of the p17/p24 Substrate Cleavage Site are Tight-Binding Inhibitors of HIV Protease. *J. Med. Chem.* **1990**, *33*, 1285–1288.
(81) Affymetrix Co. of Santa Clara, CA, a subsidiary of Affymax, uses the VLSIPS process for creation of oligonucleotide libraries on chips for the purposes of genetic typing and high-through-put DNA sequencing.